THE AMERICAN

MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 1

CONTENTS

(*Continued on inside cover*)

JANUARY 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

THIRTEEN COLORFUL VARIATIONS ON GUTHRIE'S FOUR-COLOR CONJECTURE

Dedicated to the memory of Oystein Ore

THOMAS L. SAATY, University of Pennsylvania

INTRODUCTION

After careful analysis of information regarding the origins of the four-color conjecture, Kenneth O. May [1] concludes that:

> "It was not the culmination of a series of individual efforts that flashed across the mind of Francis Guthrie while coloring a map of England . . . his brother communicated the conjecture, but not the attempted proof to De Morgan in October, 1852."

His information also reveals that De Morgan gave it some thought and communicated it to his students and to other mathematicians, giving credit to Guthrie. In 1878 the first printed reference to the conjecture, by Cayley, appeared in the Proceedings of the London Mathematical Society. He wrote asking whether the conjecture had been proved. This launched its colorful career involving a number of equivalent variations, conjectures, and false proofs, which to this day, leave the question of sufficiency wide open in spite of the fact that it is known to hold for a map of no more than 39 countries.

Our purpose here is to present a short, condensed version (with definitions) of most equivalent forms of the conjecture. In each case references are given to the original or related paper. For the sake of brevity, proofs are omitted. The reader will find a rich source of information regarding the problem in Ore's famous book [1], "The Four-Color Problem".

A number of conjectures given here are not in any of the books published so far. Others are found in some but not in others. Even though this array of conjectures may not be complete, it is hoped that the condensed presentation and its order

---

Thomas L. Saaty is Professor in the Graduate Groups in Applied Mathematics, Operations Research, and Peace Research of the University of Pennsylvania. After receiving his Ph.D. in Mathematics from Yale, under E. Hille, and spending a year at the Sorbonne in Paris, he worked with the MIT Operations Evaluations Group. Later he was Director of Advanced Planning and also Head of the Mathematics Branch of the Office of Naval Research, and Scientific Liaison Officer at the U.S. Embassy, London. In the period 1963–1969 he was involved in mathematical research on negotiation and bargaining at the Arms Control and Disarmament Agency, Washington, D. C. From 1965 to 1967 he served as Executive Director of the Conference Board of the Mathematical Sciences. He is the author, co-author, or editor of 13 books in mathematics, operations research, arms control, and compact cities of the future. His areas of interest are graph theory, optimization, and applications to social problems. He is a member of the Institute for Strategic Studies, London, and was recently elected to membership in the Academy of Sciences of Spain. *Editor*.

would givé the interested reader a feeling of the depth and variety in which the problem has been examined by a large number of people.

We have intentionally avoided extending the concepts to important areas of graph theory which do not have direct bearing on the conjectures given here. Otherwise, there would be no end to this paper.

CHAPTER I: THEME

1. Basic definitions and statement of the conjecture.

1.1 DEFINITION: A **graph** is a triple (V, E, Φ) where V is a finite nonempty set called the set of **vertices,** E a (possibly empty) finite set called the set of **edges,** with $E \cap V = \emptyset$, and $\Phi\colon E \to V \,\&\, V$ is a function called the **incidence mapping.** Here $V \,\&\, V$ is the unordered product of V with itself; i.e., if $(u \,\&\, v) \in V \,\&\, V$ then $(u \,\&\, v) = (v \,\&\, u)$. If $\Phi(e) = (v \,\&\, w)$, then we say that v and w are **incident** with e. Two vertices connected by an edge (incident with the same edge) are said to be **adjacent.** They are called the **end points** of the edge. Two edges with a vertex in common are also called adjacent.

A graph is **simple** if it has no loops or parallel edges. (An edge is a **loop** if both of its end points coincide; two edges are **parallel** if they have the same end points.)

1.2 DEFINITION: A sequence of n edges $e_1, \cdots, e_n$ in a graph G is called an **edge progression** of length n if there exists an appropriate sequence of $n + 1$ (not necessarily distinct) vertices $v_0, v_1, \cdots, v_n$ such that e_i is incident with v_{i-1} and v_i, $i = 1, \cdots, n$. The edge progression is **closed** (open) if $v_0 = v_n$ $(v_0 \neq v_n)$. If $e_i \neq e_j$ for all i and j, $i \neq j$, the edge progression is called a **chain progression.** The set of edges is said to form a **chain.** The chain is a **circuit** if $v_0 = v_n$. If the vertices are also distinct, we have a **simple chain progression,** the edges form a simple chain. In this case, if only $v_0 = v_n$ and all other vertices are distinct, the edges are said to form a **simple circuit.** The **length** of (number of edges in) a longest simple circuit is called the **circumference** of G. Frequently one abbreviates a "simple circuit" by a "circuit".

1.3 DEFINITION: The **degree** (or valence) of a vertex is the number of edges incident with that vertex.

1.4 DEFINITION: A graph is: **planar** if it can be embedded (drawn) in a plane (or on a 2-sphere) such that no two edges meet except at a vertex; **connected** if each pair of vertices can be joined by a chain; **complete** if each vertex is connected by an edge to every other vertex; k**-partite** if its vertices can be partitioned into k disjoint sets so that no two vertices within the same set are adjacent; and **complete** k**-partite** if every pair of vertices in different sets are adjacent. A **connected component** of a graph is a maximal connected subgraph.

Note that a graph is bipartite if and only if every circuit has even length. (Bipartite means 2-partite.)

1.5 DEFINITION: A **map,** or **planar map,** M consists of a planar graph G together with a particular drawing, or embedding, of G in the plane. We call G the **underlying graph** of M and write $G = U(M)$. The map M divides the plane into connected components which we call the **regions,** or **faces,** or **countries,** of the map. Two regions are **adjacent** if their boundaries have at least one common edge, not merely a common vertex. We refer to the edges in the boundary of a region as its **sides.**

Note that a graph may be embedded in the plane to produce several different maps. For example, the graph which consists of a square and two triangles all meeting at one vertex may be embedded in the plane in several ways—one has both triangles on the inside of the square, another has one triangle inside and one triangle outside the square. In the second map there is no four-sided region, while in the first map the region exterior to the square has four sides.

1.6 DEFINITION: A **k-coloring** of a map (sometimes called a **proper k-coloring**) is an assignment of k colors to the countries of the map in such a way that no two adjacent countries receive the same color. A map is **k-colorable** if it has a k-coloring.

1.7 CONJECTURE C_0: *Each planar map is 4-colorable.*

K. May points out that the four-color conjecture belongs uniquely to Francis Guthrie and could fairly be called "Guthrie's Conjecture". That four colors are necessary can be seen from the two figures below, the first of which has four regions, each of which is adjacent to the remaining three. However, this type of condition need not hold in order that four colors be necessary as illustrated by the second figure.

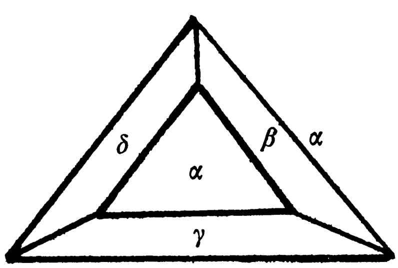


FIG. 1

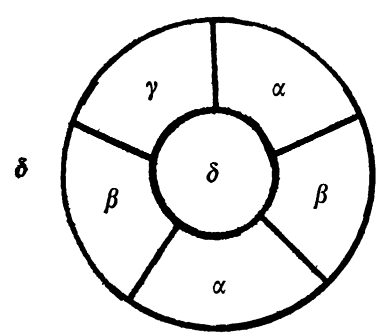


FIG. 2

2. Historical highlights. Because of the many valuable contributions of many people to the four-color problem, we are reluctant to appear to give special mention to some contributors but not to others. Nevertheless, we thought it would be useful to give a brief summary of some of the historical events relating to the conjecture and some of its variations. Occasionally it is difficult to pinpoint the exact date of an idea. The best one can do is give the year of its appearance in print. The names

of G. A. Dirac and W. T. Tutte may well be added here for their many contributions to ideas occurring in the context of the four-color problem.

1852. F. Guthrie [May 1] communicated the four-color conjecture to De Morgan.

1878. A. Cayley [1] published an inquiry as to whether the conjecture had been proven.

1879. A. B. Kempe [1] published a "proof" of the conjecture. W. E. Story [1] used Kempe's work to show that the conjecture for arbitrary planar maps can be reduced to cubic maps.

1880. P. G. Tait [1] reduced the conjecture to the colorability of the edges of cubic maps.

1890. P. J. Heawood [1] pointed out the error in Kempe's proof and salvaged enough to prove the sufficiency of 5 colors for planar maps.

1891. J. Petersen [1, p. 219] proved that either the vertices of a planar cubic map can be toured by a Hamiltonian circuit or by a collection of mutually exclusive subcircuits.

1912. O. Veblen [1] transformed the conjecture into equivalent assertions in projective geometry and the solution of simultaneous equations. G. D. Birkhoff [1] introduced a version of chromatic polynomials.

1922. P. Franklin [1] showed that a map with 25 or fewer regions is 4-colorable.

1925. A. Errera [1], referring to Franklin's result that a map requiring five colors must have at least 26 regions, proved that such a map must include at least 13 pentagons.

1926. C. N. Reynolds [1] showed that a map with 27 or fewer regions is 4-colorable.

1931. H. Whitney [1] used the notion of the dual graph and proved that the dual graph to a loopless cubic map always has a Hamiltonian circuit. He also proved the equivalence of the four-color conjecture and the fact that if a planar graph is Hamiltonian, it is 4-colorable.

1932. H. Whitney [4] studied chromatic polynomials.

1936. D. König [1] published the first book on graph theory with notions later used to formulate conjectures equivalent to the four-color problem.

1937. C. E. Winn [1], considering Franklin's paper which was to be published in 1938, in which Franklin proved that a map which requires five colors must have at least 32 regions, showed that it must contain at least 2 regions bounded by more than six edges (see Ball and Coxeter [1, p. 230]).

1938. P. Franklin [2] extended the number to 31 regions (thus if a map were to require 5 colors, it must have at least 32 regions). He also showed that such a map must include at least 15 pentagons.

1940. C. E. Winn [4] extended the number of regions in a 4-colorable map to 35.

1941. R. L. Brooks [1] proved an important theorem giving a bound on the chromatic number of a graph.

1943. H. Hadwiger [1] gave his well-known conjecture of which the four-color problem is a special case.

1952. Dynkin and Uspenskii [1] first published a small book of elementary exercises on the coloring problem.

1959. G. Ringel [1] published the first major book on the coloring of maps and graphs.

1967. O. Ore [1] published the now classic book on the subject containing a number of new ideas.

1969. O. Ore and G. J. Stemple [1] increased the number of regions to 39.

Several other books now include chapters on the theory of graphs and on coloring problems. The leading texts fully given to the subject are the books by C. Berge [1], F. Harary [2], B. Roy [1], and by W. T. Tutte [11]. No library is complete without them. One may also refer to Busacker and Saaty [1], Franklin [3], and Liu [1].

CHAPTER II: VARIATIONS ON THE THEME

1. Duality and coloring. Given a map M there is another graph $D(M)$ which we can derive from it. Replace each region by a vertex, or **capital,** and join two capitals by as many parallel edges as there are edges common to the boundaries of both corresponding regions. Thus an edge which lies on the boundary of only one region in M produces a loop in $D(M)$.

1.1 DEFINITION: The graph described above is called the **dual graph** $D(M)$ of the map M.

Note that the dual graph is the underlying graph of a (dual) map.

1.2 DEFINITION: A k-**coloring** (or **proper** k-**coloring**) of a graph is an assignment of k colors to the vertices of the graph in such a way that no two adjacent vertices receive the same color. A graph is k-**colorable** if it has a k-coloring.

Thus, a map is k-colorable if and only if its dual graph $D(M)$ is k-colorable.

1.3 PROPOSITION: *Let M be any map. We may subdivide the edges of $U(M)$—i.e., introduce vertices of degree 2—to obtain a new map M' for which $U(M')$ is simple. Hence, to 4-color M, it suffices to 4-color M'.*

Thus, in coloring a map M we may always assume that $U(M)$ is simple. Note, however, that by making $U(M)$ simple we may force $D(M)$ to be non-simple. For example, if $U(M)$ consists of a loop, $D(M)$ is a simple edge. Subdividing $U(M)$ introduces parallel edges in $D(M)$.

If G is a graph, we write $S(G)$ for the simple graph obtained from G by deleting loops and replacing parallel edges by a single edge. Obviously, we have the following result:

1.4 LEMMA. *G is k-colorable if and only if $S(G)$ is k-colorable.*

1.5 CONJECTURE C_1: *Every planar graph is* 4-*colorable.*

REMARK: Some misunderstanding can result from not making the distinction between Conjectures C_0 and C_1. Sometimes authors speak of graphs in both cases and refer to coloring regions or vertices as the case may be. Perhaps it is best when using Conjecture C_0 to refer to a map and when using Conjecture C_1 to refer to a graph, the first suggesting a coloring of regions and the second a coloring of vertices. Thus, in the sequel when speaking of equivalent conjectures, whenever we speak of graphs, the equivalence is to Conjecture C_1. The equivalence of Conjecture C_1 and Conjecture C_0 follows from the definition of a dual graph. Characterization of planar graphs in terms of an *abstract* duality was established by H. Whitney **[1]**. In particular he showed that if M is planar, so is $D(M)$.

As a consequence of the easier half of the theorem of Kuratowski **[1]**, proving that the complete graph on five vertices is nonplanar, one can conclude that there are no planar maps in which five countries are pairwise adjacent.

Heawood's proof that any planar map can be 5-colored is inductive and surprisingly simple, and it exemplifies the many ingenious approaches which have been taken in pursuit of the four-color problem. However, rather than prove the sufficiency of five colors, we prefer to use the method of Heawood's proof to show that a planar map containing a region with no more than four sides must be 4-colorable, provided that we first assume it is *irreducible* — i.e., minimally non-4-colorable. Thus, we shall see that every region in an irreducible map has 5 or more sides.

Note that in particular any map with at most 12 regions has some region with no more than four sides. To see this, suppose that the map has n vertices, m edges, and r regions. Then Euler's formula (satisfied by planar maps) gives

$$n - m + r = 2. \tag{1}$$

Assuming without loss of generality that M has no vertices of degree one or two, we always have $3n \leqq 2m$ and if we assume that every region of a 4-colorable map is bounded by at least five edges, then $5r \leqq 2m$. Substitution in (1) gives $m \geqq 30$. Substituting $3n \leqq 2m$ alone in (1) gives $m \leqq 3r - 6$ which for $r < 12$ gives $m < 30$. Hence, a map of less than 12 regions has at least one region bounded by less than five edges.

To color the vertices of the dual graph $D(M)$ of such a map M with four colors, let v be the vertex adjacent to (1) four other vertices, v_1, v_2, v_3, v_4, or (2) three other vertices (the proof of this case is trivial).

By minimality of M, we may assume that on suppressing v and its four incident edges, the vertices of the resulting graph have been colored with four colors, which we denote by c_1, c_2, c_3, c_4. Let this assignment result in giving v_i the color c_i, $i = 1, \cdots, 4$. See Fig. 3.

Now if there is a chain from v_1 to v_3 whose vertices are alternately colored with c_1 and c_3 starting at v_1 and ending at v_3, then there cannot be a chain whose vertices

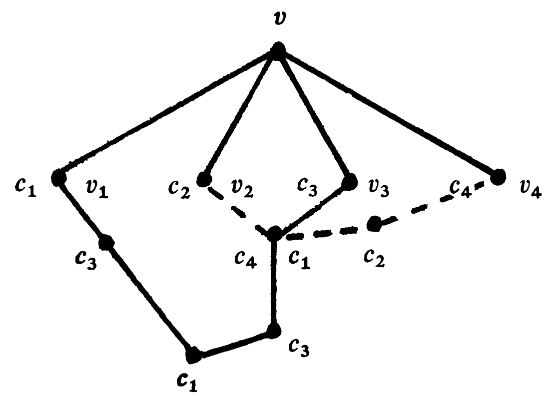


FIG. 3

are alternately colored with c_2 and c_4 starting at v_2 and ending at v_4. Otherwise the two chains must cross (see diagram) at a vertex whose color would conflict in the two chains. Thus, the second chain of alternating colors may have the colors of its vertices reversed. In that case, v_2 could be assigned the color c_4, and the remaining color c_2 would then be assigned to v.

If the first chain starting at v_1 does not terminate at v_3, then the color of its vertices may be reversed, assigning c_3 to v_1, leaving c_1 to be assigned to v. This completes the argument.

In every planar map there is at least one region bounded by five or fewer edges. Otherwise we have $3n \leqq 2m$, $6r \leqq 2m$, and substitution in Euler's formula gives $2m/3 - m + 2m/6 \geqq 2$, a contradiction.

A slight adaptation of the foregoing approach, again applied inductively to a vertex of the dual graph which has five or less neighbors, can be used to prove the following theorem (Heawood [**1**]).

1.6 THEOREM. *Any planar graph is 5-colorable.*

Of course, the problem is to show that any planar graph is 4-colorable.

Sketch of Heawood's Argument (Fig. 4). Heawood's counterexample [**1**] is directed at Kempe's chain coloring reversals. He is not concerned with whether one can by a judicious choice recolor some of the vertices. The above example with 25 vertices is known to be 4-colorable by existing theory.

Using the inductive argument on the number of vertices n, assume that every planar graph on $(n - 1)$ vertices is 4-colorable. Consider a graph on n vertices and remove a vertex v (which has five neighbors) and its connecting edges and 4-color the resulting graph on $n - 1$ vertices. Suppose the coloring is as shown. Reinstate v and attempt to color the resulting graph.

There is a b–g chain from 2 to 4. There is also a b–y chain from 2 to 5.

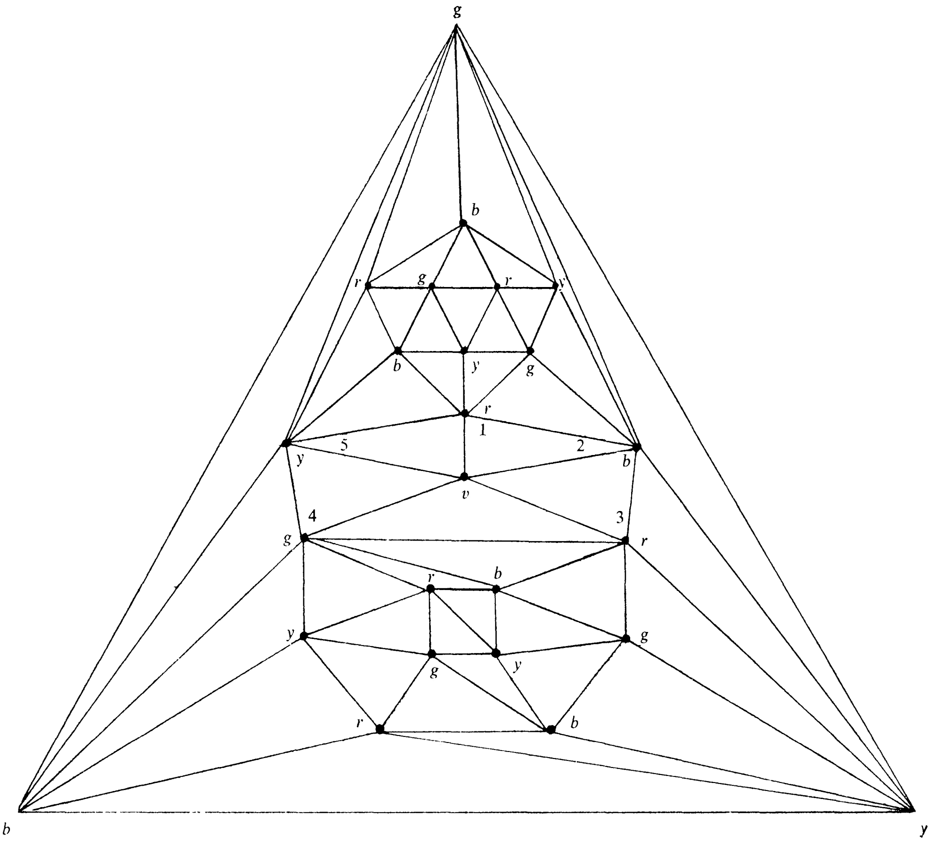


FIG. 4 — Heawood's counterexample to Kempe's proof

Reversal of colors on either chain will not free a color for v. This leaves r in two places. Now there is no r–g chain from 1 to 4. Therefore, one can reverse r to g in the r–g chain starting at 1. But the other r at 3 must also be turned to g or to y to obtain a spare color for v. This is not possible because 4 which has color g is adjacent to 3 which will become colored with g. On the other hand, if we reverse colors on the r–y chain starting at 3, the two vertices of the outer triangle which are connected by an edge would both be assigned r by the r–g and r–y reversals, starting at 1 and at 3 respectively, contradicting proper coloring. Thus, one cannot replace r by g at both 1 and 3 nor by g at 1 and by y at 3. Note that at 1, r cannot be turned to y because it is adjacent to a y at 5. Heawood [1] wrote "Unfortunately, it is conceivable that though either transposition would remove an r both may not remove both r's." (It is clear that reversal of colors on the y–r chain starting at 5 followed by a reversal on the r–g chain starting at 1 frees the color y for v, but this does not justify Kempe's argument.) See also Saaty [1].

2. Cubic Maps.

2.1 DEFINITION: A graph is **cubic (normal, regular, regular of degree three, trivalent)** if all of its vertices are of degree 3. A map is **cubic (normal, regular, trivalent)** if $U(M)$ is cubic.

2.2 DEFINITION: A graph is **bridgeless** (or doubly edge-connected) if there is no edge whose removal disconnects the vertices (i.e., after any edge is removed, it is still possible to connect any two vertices by a chain). An edge e is called a **bridge** (or **isthmus**) if the set of vertices can be partitioned into two sets T and U such that e is the only edge with one end point in T and the other end point in U.

Obviously, a graph is bridgeless if and only if it has no bridges. A map M is bridgeless if $U(M)$ is bridgeless.

REMARK: In a cubic graph, a loop is counted twice. A cubic graph with a loop must have a bridge.

In coloring maps, we can really assume that the maps are bridgeless as the following argument will show.

2.3 LEMMA. *Let e be an edge of a graph G. Then e is a bridge if and only if e lies on no circuit.*

2.4 LEMMA: *Let e be an edge of a map M (i.e., e is an edge of U(M)). Then e is a bridge if and only if e lies on the boundary of exactly one region.*

2.5 THEOREM. *Let M be any map. Then there exists a map M′ such that* (i) *M′ is bridgeless,* (ii) *M′ can be k-colored if and only if M can be k-colored.*

Proofs: Lemmas 2.3 and 2.4 are trivial. We obtain M' by simply shrinking each bridge to a point. By Lemma 2.4, M' satisfies the conclusions of the theorem.

2.6 CONJECTURE C_2: *Every bridgeless cubic planar map is* 4*-colorable.*

As we indicated in Section 2 of Chapter 1, reduction of the four-color problem to cubic maps is due to Story. A proof of the equivalence of Conjectures C_0 and C_2 is given in Harary's book [**2**, p. 132]. To go from any map to a cubic map, each vertex is blown into a polygon with as many vertices as there are edges incident with the vertex. Out of each of these vertices of the polygon emanates one of the edges. Thus, each vertex is of degree three, and the resulting map is cubic. After coloring the cubic map, the added polygons are contracted back to the vertex to obtain a coloring for the original map.

2.7 DEFINITION: A region is called **odd** (even) if it is bounded by an odd (even) number of edges. A circuit is called odd (even) if its length is odd (even).

REMARK: In Problem E 1756, this MONTHLY, 72 (1965) p. 76, it is shown that in a 4-colored cubic map, the number of odd regions colored by any two colors is even.

2.8 DEFINITION: A map, all of whose vertices have even degree, is said to be **triangle-colored** when its regions can be colored in two colors such that all regions colored with one of the colors are triangles.

2.9 CONJECTURE C_3: *The vertices of a planar triangle-colored map without multiple edges and all of whose vertices have degree four can be* 3-*colored.*

This conjecture is equivalent to Conjecture C_1 (Ore [**1**, p. 126]).

2.10 DEFINITION: We call a map M **triangular** if its dual $D(M)$ is a cubic graph. We shall discuss triangular maps later.

3. Edge Coloring.

3.1 DEFINITION: A (proper) coloring of the edges of a cubic map (called a **Tait-coloring or edge-coloring**) is a 3-coloring of the edges such that all three edges incident with the same vertex have different colors.

3.2 CONJECTURE C_4: *The edges of a bridgeless cubic planar map are* 3-*colorable.*

The equivalence of Conjectures C_2 and C_4 is due to P. Tait [**1**]. Proofs are found in Ball and Coxeter [**1**, p. 226], Ore [**1**, p. 121], and Liu [**1**, p. 253] (in dual form—see Conjecture C_5). A cubic map with a bridge has no Tait coloring. According to a previous remark, if the map has a loop, it has no Tait coloring.

3.3 CONJECTURE C_5: *The edges of a triangular map can be colored with three colors so that the edges bounding every triangle are colored distinctly.*

Let us actually see how to construct Tait-colorings from region-colorings and region-colorings from Tait-colorings.

Suppose that we are given a bridgeless cubic map M whose regions have been 4-colored using colors 0, 1, 2, 3. We may then Tait-color the edges according to the following scheme:

Color edge:	if edge lies on boundaries of	regions colored:
α		0 and 1, or 2 and 3
β		0 and 2, or 1 and 3
γ		1 and 2, or 0 and 3

It is easy to check that this scheme actually works.

Conversely, suppose we are given a Tait-coloring of the edges of M using the colors α, β, γ. Those edges labelled α and β form disjoint simple circuits (of even length) which we call α-β circuits.

Now every region R of M is contained in the interiors of either an odd or an even number of α-β circuits. Let us pre-color R with $1'$ if R is contained in an odd number of α-β circuits and $0'$ if R is contained in an even number of α-β circuits. Similarly, we have α-γ circuits and every region R of M is contained in either an even or odd number of α-γ circuits. In the former case, we pre-color R with $0''$ and in the latter case with $2''$. Now color the regions of M according to the following scheme:

Color region:	if region has already been	pre-colored:
0		$0'$ and $0''$
1		$1'$ and $0''$
2		$0'$ and $2''$
3		$1'$ and $2''$

Thus, each region is pre-colored twice and two regions are colored the same if and only if both of their pre-colorings are the same.

This yields a proper coloring of the regions. For if two regions R_1 and R_2 have a common edge e, then e may be colored either α, β, or γ. If e is colored β, then e lies on exactly one α-β circuit C which contains either R_1 or R_2, but not both, in its interior. Hence, R_1 and R_2 are pre-colored with $1'$ and $0'$ or $0'$ and $1'$, respectively. Thus, they cannot be colored the same. The same argument holds when e is colored γ. If e is colored α, then e lies on both an α-β and an α-γ circuit so the argument above shows that both pre-colorings of R_1 and R_2 are different, and we may again conclude that R_1 and R_2 are colored differently.

3.4. DEFINITION: The **line** or **interchange** graph $L(G)$ of a given graph G (without multiple edges) is obtained by associating a vertex with each edge of the graph and connecting two vertices by an edge if and only if the corresponding edges of the given graph are adjacent.

3.5 CONJECTURE C_6: *The vertices of the line graph of a bridgeless cubic planar map can be colored with* 3 *colors.*

The equivalence of Conjectures C_4 and C_6 is trivial.

For more information on line graphs, see Ore [**1**, p. 124]. Ore quotes the following two results of Sedláček [**1**]:

3.6 THEOREM. *A planar graph G has a planar line graph $L(G)$ if and only if no vertex in G has degree exceeding* 4, *and when a vertex has degree* 4, *then its removal must disconnect the graph.*

3.7 THEOREM. *If G is nonplanar, then $L(G)$ is nonplanar.*

4. Hamiltonian circuits.

4.1 DEFINITION: A graph is said to be **Hamiltonian** if it has a simple circuit called a **Hamiltonian circuit** which passes through each vertex exactly once.

It is clear that if M is a cubic map and $U(M)$ has a Hamiltonian circuit C, then the edges of the map M can be 3-colored. (Recall that in the cubic graph $U(M)$, there must be an even number of vertices because $3n = 2m$ where m is the number of edges. Thus two colors are alternately assigned to the edges of C, and the third color is assigned to the remaining edges.) This implies that M is 4-colorable.

4.2 CONJECTURE C_7: *Every Hamiltonian planar graph is* 4-*colorable.*

Proof of the equivalence of Conjectures C_1 and C_7 is due to Whitney [**1**]. It is clear that if a planar graph is 4-colorable, then also every Hamiltonian planar graph is 4-colorable. The proof of the converse is not obvious. It depends on the result of Whitney [**1**] that every maximal planar graph (see 6.4) has a Hamiltonian circuit.

4.3 CONJECTURE C_8: *It is possible to* 4-*color the vertices of a planar graph consisting of a regular polygon of n sides with non-crossing diagonals dividing the interior of the polygon into triangles and with non-crossing edges dividing the exterior of the polygon into triangles.*

Whitney [**1**] proves the equivalence of Conjecture C_8 and Conjecture C_0. Conjecture C_8 is essentially Conjecture C_7. For a discussion of the following conjecture, see Ball and Coxeter [**1**, p. 226], Petersen's 1891 paper, page 219, and Ore [**1**, p. 121].

4.4 Conjecture C_9: *In a bridgeless cubic map it is possible either to tour all the vertices by a Hamiltonian circuit or to make a group of mutually exclusive subcircuits (subtours) of the vertices in several even-length simple circuits.*

The equivalence of this conjecture with Conjecture C_4 is essentially due to Tait who preceded Petersen and is easy to establish. We give a sketch here. We assume that the edges have been 3-colored. We start at any vertex and follow a chain whose edges alternate with two colors. Such a chain must return to its starting point to form a simple circuit. The reason is that since the degree of each vertex is 3, and the three edges meeting at any vertex have all three colors, returning to an intermediate vertex would mean that the tour would have used the third color contrary

to assumption. Because of connectedness, the tour must return to the starting vertex and hence it must have even length. If all the vertices are included in this tour, we have a Hamiltonian circuit of even length. Otherwise, the process is repeated on the remaining vertices to form another simple circuit (subtour) disjoint from the first and so on.

If, on the other hand, we have the disjoint subtours of even length, we color their edges alternately with two colors and assign the third color to edges not on any subtour. In this manner we can 3-color the edges.

4.5 Definition: Let G be a graph and G' a subgraph. We call G' a **section graph** of G if two vertices are adjacent in G' whenever they are adjacent in G.

Thus, a section graph of G is determined by its set of vertices. Let G be a graph which has been 4-colored (say red, blue, yellow, and green).

4.6 Definition: A **Kempe chain** in G is a connected component of a section graph determined by all of the vertices in two of the colors.

4.7 Definition: Let M be a map which has been 4-colored. Then a collection of regions in M forms a **Kempe chain** in M if its dual is a Kempe chain in (DM).

4.8 Definition: A family of disjoint simple closed curves of even length including every vertex in M is called a **Tait cycle.**

Suppose we have a red-blue Kempe chain K in a cubic map M. Let R_1 be a region of K. If R_2 is a region not in K and R_1 and R_2 are adjacent, then R_2 must be colored either yellow or green. Thus every edge on the boundary of K separates a red or blue face from a yellow or green face, and hence, by our construction scheme, we can Tait-color the edges of M using only two colors for the boundary edges of K. This implies that the boundary of K consists of a family of even-length simple closed curves. Moreover, since every vertex in M is on the boundary of three differently colored faces, every vertex belongs to one, and only one, of the simple closed curves in the boundary of a Kempe chain. Thus a 4-colored cubic map has a Tait cycle. Note that in fact the coloring has three Tait cycles, one for each separation of the four colors into pairs.

One can reformulate Conjecture C_9 in terms of Tait cycles. We use this nomenclature later on in the paper.

4.9 Definition: A graph is said to be **p-connected** if each pair of vertices v and w is connected by at least p chains which have no vertices in common other than v and w.

A graph G is p-connected if and only if G is not disconnected or made trivial by the removal of $p - 1$ or fewer vertices.

There are special types of graphs which are known to be Hamiltonian; e. g., complete graphs with $n \geqq 3$ vertices. As another example, Tutte [3] has proved that

a 4-connected planar graph with at least two edges has a Hamiltonian circuit. Whitney [1] has shown that if M is a cubic map then $D(M)$ has a Hamiltonian circuit.

That not every planar graph is Hamiltonian is illustrated in Fig. 5 which shows a graph with 20 vertices and 12 pentagonal faces. It is easy to show that this graph is 4-colorable.

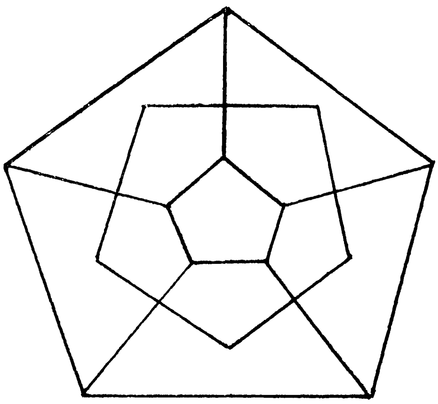
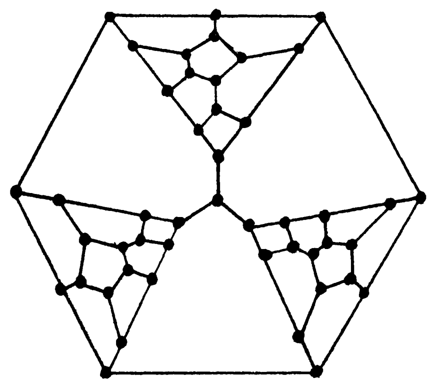


FIG. 5



Tutte's counterexample
FIG. 6

Dirac [3] has shown that each graph on n vertices, the degree of each vertex of which is at least $n/2$, has a Hamiltonian circuit. L. Posa [1] proved that a graph on $n \geqq 3$ vertices has a Hamiltonian circuit if for each integer i with $1 \leqq i < n/2$, the number of vertices of degree not exceeding i is less than i. See the book by B. Roy [1] for additional results.

Tait [3] once conjectured that every 3-connected planar graph is Hamiltonian but Tutte [3] gave a counterexample (Fig. 6) with 46 vertices. Had Tait's conjecture been true, the truth of Conjecture C_0 would have followed. For as we shall see in the last chapter, to prove Conjecture C_0, it suffices to show that every cubic map M with $U(M)$ 3-connected can be 4-colored. But Tait's conjecture would imply that every such map had a Hamiltonian circuit and hence was 4-colorable.

Tait himself did not supply an adequate proof as to how the four-color conjecture would be true if his conjecture were true. He thought his conjecture was true from all the evidence he had. Chuard [1] went on to "complete" the story in 1932. Doubts as to the validity of Chuard's claim were expressed by Pannwitz [1].

In any event, Tutte's example has made the entire debate academic as a means of settling the four-color conjecture.

5. Flow ratio.

5.1 DEFINITION: A graph is called **directed** or **oriented** if each edge is assigned a direction (indicated by an arrow) from one of its end vertices toward the other.

5.2 DEFINITION: The flow ratio of a simple circuit is the ratio m_1/m_2, where m_1 and m_2 are the numbers of edges of the circuit directed clockwise and counter-clockwise around the circuit with $m_1 \geqq m_2$. If $m_1 < m_2$ then the roles of m_1 and m_2 are interchanged (the flow ratio may be $+\infty$).

5.3 CONJECTURE C_{10}: *The edges of a planar graph can be oriented in such a way that the flow ratio of each cycle is at most* 3.

A proof of the equivalence of Conjectures C_1 and C_{10} is due to Minty [**1**]. Actually Minty proves the equivalence of k-colorability to the fact that the flow ratio of each circuit does not exceed $k - 1$.

5.4 CONJECTURE C_{11}: *The edges of a planar graph can be so directed that for any circuit C with $m(C)$ edges and any direction associated with the circuit (clockwise or counter-clockwise), the number of edges of C oriented opposite to the given direction and denoted by $m_1(C)$ satisfies*

$$m_1(C) \geqq \tfrac{1}{4}m(C).$$

This is obviously equivalent to the previous result (see Ore [**1**, p. 104]).

6. Partition of vertices; chromatic number. When the vertices of a planar graph are 4-colored, they are divided into four disjoint sets such that the vertices in each set are assigned the same color and no two vertices of the same color are joined by an edge. Clearly a graph can be 4-colored if and only if it is 4-partite. Each pair of these four sets, together with their interconnecting edges, forms a bipartite graph.

6.1 DEFINITION: A planar graph is said to have **bipartite dichotomy** if there is a disjoint decomposition of its vertices into two sets such that each set defines a bipartite graph.

We sometimes call a bridge a **separating edge.**

6.2 CONJECTURE C_{12}: *The dual of a planar map without separating edges has a bipartite dichotomy.*

6.3 CONJECTURE C_{13}: *Any planar graph without loops has a bipartite dichotomy.*

See Ore [**1**, page 105] for the equivalence of these conjectures to Conjecture C_1.

6.4 DEFINITION: A graph G is called **maximal planar** if it is planar and has no loops and no multiple edges and it is not possible to add a new edge to G without violating one of these restrictions.

REMARK. The following statements are equivalent:

(i) G is maximal planar;

(ii) For every map M with $G = U(M)$, M is triangular;

(iii) There exists a triangular map M with $G = U(M)$.

It is known that every uniquely 4-colorable planar graph is maximal planar (Harary [**2**, page 140]).

6.5 CONJECTURE C_{14}: *Every maximal planar map has a bipartite dichotomy.*

The equivalence of Conjectures C_1 and C_{14} is proved in Ore [**1**, page 122].

6.6 DEFINITION: The **chromatic number** $\chi(G)$ of a graph G is the minimum number of disjoint subsets into which its vertices can be partitioned such that no two vertices in the same subset are adjacent.

6.7 CONJECTURE C_{15}: *The dual graph G of a planar map satisfies* $\chi(G) \leqq 4$.

REMARK. Ershov and Kozhukhin [1] have shown that a connected graph G with n vertices and m edges satisfies the following bounds on its chromatic number (using $[x]$ and $\{x\}$ to denote the integral and fractional parts of x, respectively):

$$-\left[-\frac{n}{[(n^2-2m)/n]}\left(1-\frac{\{(n^2-2m)/n\}}{1+[(n^2-2m)/n]}\right)\right] \leqq \chi(G) \leqq \left[\frac{3+\sqrt{9+8(m-n)}}{2}\right].$$

If the vertices of a graph G are numbered $i = 1, \cdots, n$ according to the decreasing order of their degree d_i, and if k is the last number of a vertex which satisfies

$$k \leqq d_k + 1, \text{ then } \chi(G) \leqq k.$$

It follows from this that $\chi(G)$ is at most equal to the highest degree of any vertex plus unity. Welsh and Powell [**1**] give an algorithm for coloring the vertices of a graph with a number of colors equal to the bound k.

6.8 DEFINITION: A graph G is called **critical**, or **vertex-critical,** (Dirac [**2**]) if after the removal of any vertex v and its connecting edges we have

$$\chi(G - v) < \chi(G).$$

G is k-critical if $\chi(G) = k$ (in which case, for every v, $\chi(G - v) = k - 1$). A graph is **edge-critical** if similar relations hold on removing an edge.

It is known (Ore [**1**, p. 164]) that the removal of a complete subgraph cannot separate a critical graph. Dirac [**3**] has shown that if a graph G is k-critical with $k \geqq 3$, then either G has a Hamiltonian circuit or the circumference of G is $2k - 2$. He has also proved that every k-chromatic graph contains a critical k-chromatic subgraph.

6.9 DEFINITION: The **chromatic index** $q(G)$ of a graph G is the smallest number of colors necessary to color its edges so that no two adjacent edges have the same color.

Thus $q(G) = \chi[L(G)]$ when G is simple.

6.10 DEFINITION: A **p-graph** is a graph with multiple edges between its vertices such that no two vertices are jointly incident with more than p edges.

Vizing [1] and Shannon [1] have shown that if d_m is the maximum degree of any vertex in a graph, then we have:

$$d_m \leqq q(G) \leqq \text{Min}\left(p, \left[\frac{d_m + 1}{2}\right]\right) + d_m.$$

It follows that if G has no multiple edges, $q(G)$ is either d_m or $d_m + 1$.

6.11 CONJECTURE C_{16}: *Let G be a planar bridgeless cubic graph. Then $q(G) = 3$.*

This conjecture is just a restatement of Conjecture C_4.

7. Partitions of edges; factorable graphs.

7.1 DEFINITION: A graph (or map) is k-**factorable** if its edges can be partitioned into edge disjoint subsets in such a way that in each subset any vertex meets exactly k edges of that subset. See König [**1**, pp. 155–195].

7.2 CONJECTURE C_{17}: *Every cubic bridgeless planar map is* 1-*factorable.*

This conjecture, first formulated by Tait in 1884, is obviously equivalent to Conjecture C_4. See also Harary [**2**, p. 135].

7.3 CONJECTURE C_{18}: *The dual of every connected planar map is the sum of three edge-disjoint subgraphs such that each vertex has either an even number of edges incident with it from each of the three subgraphs or it has an odd number from each of them.*

The equivalence of Conjectures C_1 and C_{18} is given in Ore [**1**, p. 103]. Alternatively, one can give a direct proof that Conjectures C_{18} and C_4 are equivalent.

8. Vertex characters.

8.1 CONJECTURE C_{19}: *It is possible to associate a coefficient $k(v)$ equal to $+1$ or -1 with each vertex in a bridgeless cubic map in such a way that $\Sigma k(v) = 0 \pmod 3$, where the summation is taken over the vertices occurring in the boundary of any region.*

Heawood [**2**] proved the equivalence of this conjecture with Conjecture C_4. A reformulation of this conjecture would be to take the above congruences and require a solution, for all of them taken together, none of whose members is congruent to zero modulo 3. Thus, if A is the (0, 1) region-vertex incidence matrix, the above is equivalent to the existence of a vector X such that $AX = 0 \pmod 3$, where none of the components of X is zero.

To see how this conjecture implies Conjecture C_4, label the edges of the map a, b, or c, such that the three edges incident with each vertex are labelled differently and the ordering of the edges $a \to b \to c$ is a clockwise rotation if $k(v) = +1$ and counter-clockwise if $k(v) = -1$. This labelling is consistent if and only if the vertex character assignment is proper; i.e., for each region $\Sigma k(v) = 0 \pmod 3$.

Using a computer code, Yamabe and Pope developed an assignment method for cubic maps of up to 36 vertices and illustrated their method by an example in their brief paper [**1**].

8.2 CONJECTURE C_{20}: *It is always possible repeatedly to cut off corners (replace a vertex by a triangle) from a convex polyhedron so that eventually a polyhedron is obtained whose faces have a number of edges which is divisible by* 3.

This conjecture due to Hadwiger [**2**] is a modification of the previous conjecture of Heawood. Cutting off corners yields vertices of degree 3, and hence the truth of the last conjecture implies Heawood's conjecture (Conjecture C_{19}). The proof in the reverse direction is more elaborate.

Conjecture C_{20} may have been suggested by a result of Heawood [**2**] in which he proved that if the regions of a map could each be subdivided (by the simple operation of adding a new edge to connect some pairs of adjacent edges thereby forming triangles) into new regions such that all the regions are bordered by edges whose number is congruent to zero mod 3, then the map is 4-colorable.

Heawood first shows constructively that such a map is 4-colorable. Then he shows that any 4-coloring of the constructed map is also a 4-coloring of the initial map by removing the edges.

9. Modular equations and Galois fields. Let $GF(k)$ denote the Galois field of order k. Thus, k is a prime power and $GF(k)$ is the unique (finite) field with k elements. Obviously, one may view a k-coloring of the vertices (or edges or regions) of a graph (or map) as an assignment of an element of $GF(k)$ to every vertex (or edge or region) of the graph (or map).

We shall consider in this section the cases $k = 2, 3, 4$. When $k = 4$, note that two elements in $GF(k)$ are equal if and only if their sum is zero. Thus, if we assign to every edge e in a bridgeless map which has been 4-colored, the sum of the colors of the two regions adjacent to e, this sum will never be zero. We may give this a matrix formulation as follows: List the edges $e_1, \cdots, e_m$ and regions $r_1, \cdots, r_n$ of a bridgeless map M. Let B be the matrix defined by putting $B_{ij} = 1$ if e_i is in the boundary of r_j and putting $B_{ij} = 0$ otherwise. Thus, each row of B contains two unit elements. B is sometimes called the **edge-region incidence matrix** of M, or simply an **incidence matrix.**

Suppose M is 4-colored. Then define a column vector $Z = (z_1, \cdots, z_n)$, where z_j is the color of the jth region, and each z_j belongs to $GF(4)$. The matrix product BZ is a column vector $P = (p_1, \cdots, p_m)$, and each p_i is the sum of two distinct elements in $GF(4)$ since e_i is on the boundary of two distinctly colored regions. Hence, each p_i is non-zero.

Now we can state the following conjecture due to O. Veblen [**1**]:

9.1 CONJECTURE C_{21}: *Let B be any edge-region incidence matrix. Then there is a*

column vector $Z = (z_1, \cdots, z_n)$ *with entries* z_j *in* $GF(4)$ *such that the matrix product* BZ *has no zero entries.*

The discussion above shows that Conjecture C_{21} is equivalent to Conjecture C_0 since the existence of the column vector Z provides us with a 4-coloring of the map.

We can also form an edge-vertex incidence matrix for a graph G and make a conjecture as before. Obviously, this procedure is equivalent to the above by duality.

We may now restate Conjecture C_{19} using the Galois field $GF(3)$. We can also define a region-vertex incidence matrix for a map M and then make the following conjecture:

9.2 CONJECTURE C_{22}: *Let* B *be the region-vertex incidence matrix of a map* M. *Then there is a column vector* $Z = (z_1, \cdots, z_n)$ *with each* z_j *in* $GF(3)$ *such that* BZ *is identically zero but no* z_j *is equal to zero.*

In an interesting generalization of these ideas, Tutte [6] has developed a framework for merging the two questions of 4-colorability and Tait-colorability of a planar map. Some of the work is motivated by a conjecture due to Tutte that any bridgeless cubic map with no Tait-coloring can be reduced to a Petersen graph (illustrated later) by deleting some edges and contracting others to single vertices. (The converse of this conjecture is known to be false—see Watkins [1]). It leads to the classification of 2-blocks where the term k-block refers to a set of points of a projective geometry $PG(q, 2)$ over the Galois field $GF(2)$ whose dimension is $\geqq k$. A k-block is **tangential** if it cannot be converted to a similar k-block by a particular process of projection. It is not known if any tangential 2-blocks (sets of points in $PG(q, 2)$ that meet every $(q-2)$ space) other than the following three exist:

— The Fano block (the plane which has exactly 7 points),

— The Desargues block (a 3-dimensional 2-block consisting of 10 points lying in three's on 10 lines in a Desargues configuration), and

— The Petersen block (this is the only 5-dimensional 2-block) which is an embedding closely related to the Petersen graph, and its existence is associated with the non-existence of a Tait coloring of the Petersen graph. In a private communication, W. T. Tutte has informed me that Mr. Biswa T. Datta of Ohio State University proved in his Ph.D. thesis that there are no 6-dimensional tangential 2-blocks.

That many excellent mathematicians have constructed erroneous proofs of the four color conjecture is perhaps a measure of the difficulty and subtlety of the problem. For example, in a recent paper, J. M. Thomas [1] attempts to prove the four color conjecture. His argument is based on Veblen's modular equation approach. However, we can point out the fault with his paper in simpler terms. Essentially, his line of argument is the slitting operation which he describes as follows:

> Let side s bond faces K, L which are unequal and do not join. Slit side s lengthwise so that its two pieces border a channel making K, L into a single face in map M' with $n-1$ faces. Let K', L' be the sums of the unknowns

at the vertices of K, L with those u, v at s omitted. A root of map system X in which u, v are numbered $+$, $-$ becomes a root of the system $X' + (K' + 1) + (L' - 1)$, where X' is the map system for M'. Conversely, such a special root of X' augmented by the values $+1$, -1 for u, v becomes a root of the map system X.

The difficulty occurs in the inductive step when he claims that he can extend a root for the slit back into a root for the original map. This means that the two regions along the slit would have to be differently colored in the slit map. If this were true, the four color conjecture would follow trivially. Unfortunately, this part of the paper appears to be as difficult as any of the other formulations.

10. Hadwiger's Conjecture.

10.1 DEFINITION: An **edge contraction** of a graph G is obtained by removing two adjacent vertices u and v and adding a new vertex w, adjacent to those vertices to which u or v was adjacent. A graph G is **contractible** to a graph H if H can be obtained from G by a sequence of edge contractions. We shall also call H a **contraction** of G. Note that G is contractible to H if and only if there is a connected homomorphism (see Ore [**2**, p. 85]) from G onto H.

10.2 HADWIGER'S CONJECTURE: *Every connected k-chromatic graph is contractible to a complete graph on k vertices.*

10.3 CONJECTURE C_{23}: *Hadwiger's conjecture is true for* $k = 5$.

The equivalence of this conjecture and Conjecture C_1 is due to K. Wagner [**4**]; a simpler proof of the equivalence has been given by R. Halin [**2**]. The truth of this conjecture for $k < 5$ was established by G. A. Dirac [**2**].

An equivalent statement of the above conjecture using the notion of conformal graphs, (Ore [**1**, p. 26]) is due to Halin [**1**].

One may use the notion of contraction to formulate a criterion for planarity which is dual to the well-known result of Kuratowski. The following theorem was discovered independently by Harary and Tutte [**1**] and by Wagner [**3**]. It was also probably known to Ringel, since he realized that any contraction of a planar graph is planar. Let K_5 denote the complete graph on 5 vertices and $K_{3,3}$ the complete bipartite graph on two sets each with three vertices.

10.4 THEOREM. *A graph is planar if and only if it has no subgraph contractible to* K_5 *or* $K_{3,3}$.

11. Amalgamation.

11.1 DEFINITION: A graph G is a **conjunction** of two disjoint graphs G_1 and G_2 if it is obtained by taking an edge $e_1 = \{a_1, b_1\}$ in G_1 and an edge $e_2 = \{a_2, b_2\}$ in G_2,

identifying (or coalescing) a_1 with a_2, deleting the edges e_1 and e_2, and introducing a new edge $e_3 = \{b_1, b_2\}$.

11.2 DEFINITION: Suppose that we are given two sets A_1 and A_2 of vertices of a simple graph G such that no edge is incident with vertices of both sets. Let μ be a 1-1 correspondence between the elements of the two sets. A **μ-coalition** of G is the graph obtained from G by identifying corresponding vertices in A_1 and A_2. Vertices which are connected by two edges as a result of the identification are connected by a single edge in the μ-coalition by eliminating one of the edges.

11.3 REMARK: Conjunctions and μ-coalitions do not decrease chromatic numbers.

11.4 DEFINITION: Let G be a conjunction of G_1 and G_2 as in 11.1. Consider a 1-1 correspondence μ between sets A_1 and A_2 where $a_1 \in A_1$, $a_2 \in A_2$, $\mu(a_1) = a_2$, and $\mu(b_1) \neq b_2$. The graph obtained by applying this μ-coalition to G is called a **merger.**

11.5 REMARK: A conjunction is a merger in which $A_1 = \{a_1\}$ and $A_2 = \{a_2\}$.

11.6 DEFINITION: A graph G is called an **amalgamation** of the disjoint graphs $G_1, \cdots, G_p$ if it is derived by repeated mergers of the G_i. A **k-amalgamation** is an amalgamation of graphs G_i, $i = 1, \cdots, p$, each of which is a complete graph on k vertices.

11.7 CONJECTURE C_{24}: *No 5-amalgamation is planar.*

The equivalence to Conjecture C_1 is given in Ore [**1**, p. 180] utilizing ideas from Hajös [**1**].

12. Other algebraic and number-theoretic approaches. The first two approaches give statements equivalent to the four color problem but for specific maps. They are useful in applying computer methods, to test whether a given map of a reasonable size (within the bounds of computer capability and of time) is 4-colorable or not. The third and fourth approaches are number-theoretic.

Diophantine Inequalities. Let the regions of a planar map be labelled $r = 1, 2, \cdots, n$. Let the variable t_r be integer-valued $0 \leqq t_r \leqq 3$. Thus, t_r assigns one of the four colors, labelled 0, 1, 2, 3 to the region whose number is r. If two regions r and s have a boundary in common, then $t_r - t_s \neq 0$. Such a relation is written down for every pair of adjacent regions. The relation for one pair may be reduced to two inequalities as follows:

$$\textit{either } t_r - t_s \geqq 1 \quad \textit{or} \quad t_s - t_r \geqq 1.$$

This pair of inequalities may now be written as

$$t_r - t_s \geqq 1 - 4\delta_{rs} \text{ and } t_s - t_r \geqq -3 + 4\delta_{rs},$$

where $\delta_{rs} = 0$ or 1. We obtain a system of such inequalities by allowing r and s to vary from 1 to n. The problem then is to determine whether it is possible to choose the integers $0 \leqq t_r \leqq 3$, $r = 1, \cdots, n$, and the binary variables $\delta_{rs}, r, s = 1, \cdots n$, such that the system of inequalities has a solution. If not, then our assumption that t_r take on only four values is untenable.

We have now proved that the following conjecture is equivalent to Conjecture C_0:

12.1 CONJECTURE C_{25}: *For every planar map the corresponding system of diophantine inequalities formulated here has a solution.*

According to G. Dantzig, this formulation was informally communicated to him by Ralph Gomory of Integer Programming fame.

Optimization. Another formulation is due to Dantzig himself [**1**, p. 549]. Referring back to Conjecture C_9, consider each subtour of a cubic map, and starting at any vertex, assign a direction to an edge. Then assign the opposite direction to the edge of the circuit adjacent to it and continue around the (even-length) circuit in this manner so that for each vertex the two edges incident with it (now called arcs) are directed away from it or directed towards it.

Label the vertices $1, 2, 3, \cdots, n$. For any pair of adjacent vertices i and j, we write $x_{ij} = 1$ if there is an arc directed from i to j. Otherwise we write $x_{ij} = 0$. Thus we always have

$$0 \leqq x_{ij} \leqq 1.$$

We also write

$$\sum_j x_{ij} = 2\delta_i \text{ where } \delta_i = 1 \text{ or } 0,$$

expressing the fact that there must be two arcs on some subtour leading away from vertex i if $\delta_i = 0$ and none if $\delta_i = 1$. The problem now is to find δ_i and x_{ij} which satisfy these three conditions. The three conditions constitute a bounded Transportation Problem, and so one may attempt to apply the techniques of integer programming to this formulation.

Arrangements. Consider the sum $a_1 + a_2 + a_3 + \cdots + a_n$. If we add brackets to this sum as one usually does to evaluate a sum, one never adds the brackets in such a way that the numbers are added more than two at a time. The result is called an **arranged sum**. For example, $a_1 + a_2 + a_3 + a_4$ can be written as an arranged sum

(1) $$((a_1 + a_2) + (a_3 + a_4))$$

or

(2) $$(((a_1 + a_2) + a_3) + a_4), \qquad \text{etc.}$$

We can define a **partial sum** to be the sum within any pair of brackets; e.g., in (2) the partial sums are

$$(a_1 + a_2), (a_1 + a_2 + a_3), (a_1 + a_2 + a_3 + a_4).$$

In (1) the partial sums are

$$(a_1 + a_2), (a_3 + a_4), (a_1 + a_2 + a_3 + a_4).$$

12.2 CONJECTURE C_{26}: *If a sum of n numbers is expressed in any two ways as an arranged sum, then one can choose integer values for the a_i's in such a way that no partial sum of either arranged sum is divisible by* 4.

For example: for (1) and (2) $a_1 = 1$, $a_2 = 1$, $a_3 = 1$, $a_4 = 2$. The equivalence of conjectures C_0 and C_{26} is due to H. Whitney [7].

Sequences.

12.3 DEFINITION: A **cartesian sequence** is a finite sequence $c(0), c(1), \cdots$ of four colors such that

(i) $c(r) \neq c(r + 1)$, $r = 0, 1, 2, \cdots$,

i.e., the same color never appears in two consecutive positions.

(ii) $c(2r) \neq c(2r + 2)$, $r = 0, 1, 2, \cdots$, is also cartesian.

12.4 CONJECTURE C_{27}: *Given any integer n and an arbitrary increasing sequence of integers* $0 \leqq i_0 < i_1 < \cdots < i_m \leqq n$, $m \leqq n$, *there exists a cartesian sequence* $c(s)$, $s = 0, 1, 2, \cdots, n$, *such that the subsequence* $d(s) = c(i_s)$ *is also cartesian*, $s = 0,1, \cdots, m$.

The equivalence of Conjectures C_0 and C_{27} is discussed by B. and R. Descartes in **[1]**.

13. Chromatic polynomials.

13.1 DEFINITION: Let $P_r(\lambda)$ be the number of ways to color an r-country map in at most λ colors. Then $P_r(\lambda)$ is called the **chromatic polynomial** of the map. It is clear that a chromatic polynomial may correspond to many maps with r countries and that a classification of r-country maps is essential in order to give $P_r(\lambda)$ more precise meaning; i.e., the number of ways to color two r-country maps can be different.

13.2 CONJECTURE C_{28}: *For any r-country planar map,* $\lambda = 4$ *is not a root of* $P_r(\lambda) = 0$.

Conjectures C_0 and C_{28} are clearly equivalent. Chromatic polynomials are due to G. D. Birkhoff **[1]** and to H. Whitney **[4]**. A chromatic polynomial is a counting method of testing the 4-colorability of a map.

In 1946 Birkhoff and Lewis **[1]** considered cubic maps (for these $P_r(0) = P_r(1) = P_r(2) = 0$) and gave the following conjecture:

13.3 CONJECTURE C_{29}:

$$(\lambda - 3)^r \ll \frac{P_{r+3}(\lambda)}{\lambda(\lambda - 1)(\lambda - 2)} \ll (\lambda - 2)^2 \quad \text{for } \lambda \geqq 4.$$

They were only able to show this for $0 \leqq r \leqq 8$. The double inequality has the following meaning: If $f(\lambda)$ and $g(\lambda)$ are polynomials, then $f(\lambda) \ll g(\lambda)$ if and only if the coefficients of $f(\lambda)$ are non-negative and not greater than the corresponding coefficients of $g(\lambda)$. Such a relation with an additional condition such as $\lambda \geqq 4$ means that the relation holds with λ replaced by $\lambda - 4$. Note that Conjecture C_{29} implies Conjecture C_{28}. Thus, Conjecture C_{29} is a strong form of Conjecture C_0.

Rota **[1]** has proved that the coefficients of every chromatic polynomial alternate in sign. Read **[2]** conjectured that in their absolute values, these coefficients strictly increase and then strictly decrease.

We now give some interesting results due to W. T. Tutte **[8]** and **[10]** on chromatic polynomials. Let M be a triangular map with k vertices. Then the chromatic polynomial of M, $P(M, \lambda)$, with respect to vertex-coloring satisfies the relation

$$|P(M, 1 + \tau)| \leqq \tau^{5-k},$$

where $\tau = (1 + \sqrt{5})/2 = 1.618$, the "golden ratio" which is one of the solutions of the quadratic equation

$$x^2 = x + 1.$$

Tutte gives this result as a theoretical explanation of the empirical observation that $P(M, \lambda)$ appears to have a zero near $\lambda = 1 + \tau$. Note that there are no λ-colorings for the case where an edge forms a loop. For any loopless triangular map T, Tutte **[10]** shows that $P(T, \tau + 2) > 0$. Since $\tau + 2 = 3.618$, this result tells something of the behavior of $P(T, \lambda)$ near $\lambda = 4$. It is known that $P(T, \lambda)$ is not positive throughout the interval $\tau + 2 < \lambda < 4$.

If the map consists of triangles except for one region which is an m-gon with $2 \leqq m \leqq 5$, then

$$|P(M, 1 + \tau)| \leqq \tau^3 + m - k.$$

Recently, Tutte **[12]** has shown that if M is a triangular map with n vertices, then

$$P(M, \tau + 2) = (\tau + 2)\tau^{3n-10}P^2(M, \tau + 1).$$

CHAPTER 3. REDUCIBILITY

1. Irreducible graphs and maps.

1.1 DEFINITION: We call a 5-chromatic planar map (graph) **irreducible** if any other planar map (graph) with fewer regions (vertices) has a chromatic number less than 5.

Thus an irreducible planar map or graph is minimal 5-chromatic.

Suppose that an irreducible map or graph exists. We shall be able to show that it **must** have certain properties which we shall call **forced**—for example, an irreducible map is forced to have simply connected regions. On the other hand, we shall show that an irreducible map may be assumed without loss of generality, to have certain **optional** properties; i.e., if an irreducible map exists, then we may construct an irreducible map possessing the optional property. For example, if an irreducible map exists, then we may construct an irreducible cubic map from it.

1.2 CONJECTURE C_{30}: *There are no irreducible graphs.*

Clearly, if Conjecture C_1 is false, then 5-chromatic planar graphs exist and, hence, so does a 5-chromatic planar graph with a minimal number of vertices. Conversely if an irreducible graph exists, then it is a 5-chromatic planar graph so Conjecture C_1 is false.

We have two main reasons for studying irreducible maps (aside from trying to show that they don't exist). First of all, in order to show that every map is 4-colorable, it suffices to show that every irreducible map is 4-colorable and hence we may assume that the map we are trying to 4-color has any forced or optional property. Secondly, we study irreducible maps in hopes of raising the Birkhoff number whose definition follows:

1.3 DEFINITION: We define the **Birkhoff number** N to be the minimum number of regions (vertices) in an irreducible map (graph).

By the usual convention, $N = \infty$ if there is no irreducible map. Any map with fewer than N regions is 4-colorable.

Very little is known about the Birkhoff number. Franklin [**1**] proved that $N \geqq 26$, and Reynolds [**1**] improved the result slightly, showing $N \geqq 28$. Franklin [**2**] improved on the improvement, obtaining $N \geqq 32$. Finally, Winn [**4**] proved that $N \geqq 36$. After a hiatus of nearly thirty years, Ore and Stemple [**1**] succeeded in raising the lower bound for N once again by proving the following theorem:

1.4 THEOREM: $N \geqq 40$.

Being irreducible is a (very!) strong requirement, and we shall be able to deduce many properties of irreducible graphs. Since loops and parallel edges do not affect colorability of a graph, we may always assume that an irreducible graph is simple. Suppose that G is a simple irreducible graph. We can embed G in a maximal planar graph $\bar{G}$ with the same number of vertices as G. $\bar{G}$ is 5-chromatic and hence irreducible. Thus, we have shown that if any irreducible planar graph exists, then there is an irreducible simple maximal planar graph. Whitney's result [**1**] guarantees that any simple maximal planar graph has a Hamiltonian circuit, and as we have seen, any map with a Hamiltonian circuit can be 4-colored. Thus, we obtain the following paradoxical result (cf. Ore [**1**, p. 193]):

1.5 THEOREM. *It is optional to assume that any map obtained by embedding an irreducible graph can be face-colored in* 4 *colors.*

Of course, this does not imply that we can *vertex*-color the graph using 4-colors. We shall see later that any triangular map except the tetrahedron is 3-colorable.

By considering maps and dualizing, we can show that the above optional conditions for irreducible graphs yield the following optional conditions for irreducible maps:

1.6 THEOREM. *The following characteristics are optional for irreducible maps*:

(a) *Bridgeless,*

(b) *Two regions meet along at most one edge,*

(c) *Cubic.*

On the other hand, certain characteristics are forced for irreducible maps. Any map divides the plane into open connected components, and the regions of the map are just the closures of these components.

1.7 THEOREM. *Let M be an irreducible map. Then any region in M is simply-connected.*

Proof: Suppose some region R is not simply-connected. Then the region divides the plane into an inside and an outside. The region R and the regions interior to it form a map M_1; the region R and regions exterior to it form a map M_2, and no internal region shares a common boundary edge with an external region. Now, since both M_1 and M_2 have fewer regions than M, we can color both M_1 and M_2 using 4 colors. By rearranging the coloring of M_1, we can insure that R receives the same color in each of the colorings of M_1 and M_2. This allows us to put the two colorings together to obtain a 4-coloring of M.

The same argument would allow us to prove that the union of any two regions in M is simply-connected. Thus, 1.6(b) is *forced.* In other words, an optional property may be forced. Actually, if Conjecture C_0 is true, *any* property is forced.

This theorem is equivalent to the fact that an irreducible planar graph has no point of articulation and thus is 2-connected (i.e., it is a block). In fact, any maximal planar, simple, irreducible graph G must be 3-connected. For if we embed G in the sphere, we obtain a triangulation so, by a theorem of Steinitz (see Steinitz and Rademacher [1], or Grünbaum [2, p. 235]) G is 3-connected. (Steinitz's theorem states that the vertices and edges of a 3-dimensional convex polyhedron constitute a planar 3-connected graph and conversely.)

Now we can use duality to prove the following theorem:

1.8 THEOREM. *Let M be an irreducible map satisfying the optional conditions* 1.6 (a), (b), (c). *Then $U(M)$ is* 3-*connected.*

Proof. Think of M as a map on the sphere. Then M is the dual of its own dual.

But the dual of M is a triangulation of the sphere and the dual of any triangulation is a convex polyhedron (see E. C. Zeeman [**1**]). Hence, M is a convex polyhedron and so, by Steinitz' theorem, $U(M)$ is 3-connected.

1.9 COROLLARY. *Let M be an irreducible map. Then it is optional that $U(M)$ be 3-connected.*

This result seems particularly interesting in view of Whitney's theorem [**8**] which says that a 3-connected planar graph embeds uniquely in the plane. Thus, Corollary 1.9 says that M is completely determined by $U(M)$. But, by the dual of Theorem 1.5, $U(M)$ can be vertex-colored in 4-colors!

2. Critical graphs and irreducibility.

2.1 DEFINITION: Let G be a graph. Then G is **contraction-critical** if any edge contraction reduces the chromatic number of G.

Obviously, any irreducible graph G is vertex-critical and contraction-critical since removing a vertex or contracting an edge both lower the total number of vertices and hence either operation decreases the chromatic number. Thus, we may examine properties of vertex-critical or contraction-critical graphs to derive information about irreducible graphs.

2.2 DEFINITION: A graph G is **k-edge connected** if removing fewer than k edges does not disconnect the graph.

Ore ([**1**, p. 165]) proves the following theorem:

2.3 THEOREM. *Any 5-chromatic vertex-critical graph is 4-edge connected.*

Analogous information about contraction-critical graphs is due to Dirac:

2.4 THEOREM. (Ore [**1**, p. 169]). *Let G be a contraction-critical graph with $\chi(G) \geqq 5$. Then G is 5-connected.*

Thus, every irreducible planar graph is 5-connected.

We can use the last result to rederive Theorem 1.5. For suppose G is irreducible planar and hence 5-connected. Tutte's theorem [**3**] (we only need 4-connected) implies that G has a Hamiltonian circuit, and we complete the argument as before. Theorem 2.4 implies that the degree of every vertex in an irreducible planar graph is at least 5. Of course, our earlier modification of Heawood's argument also proves this fact.

2.5 DEFINITION: Let G be a graph. We call a set T of vertices of G a **minimal disconnecting set** if $G - T$ is disconnected or trivial, but no proper subset of T has this property.

The preceding theorem shows that a minimal disconnecting set T must contain at least 5 vertices if $\chi(G) \geqq 5$. If T is a minimal disconnecting set in G, the section graph determined by T, $G(T)$, is called the **separating** graph. What properties must

a separating graph have? The following theorem (Ore [**1**, p. 192]) provides a partial answer.

2.6 THEOREM. *Let G be a maximal planar graph with minimal disconnecting set T. Then $G(T)$ is a simple circuit.*

2.7 THEOREM. *Let G be a contraction-critical 5-chromatic planar graph. Then G cannot be separated by a simple circuit C of length five except when one of the connected components of $G - C$ is a single vertex which is adjacent (in G) to every vertex of C.*

Let us translate this result into a statement about maps.

2.8. DEFINITION: A sequence $R_1, R_2, \cdots, R_p$ of regions in a map with R_i adjacent to R_{i+1}, $1 \leqq i \leqq p - 1$, R_p adjacent to R_1, and no other pairs R_i and R_j adjacent is lcalled a **ring** of length p, or p-**ring.**

Obviously, a ring of length p in a map M corresponds to a simple circuit of length p in $D(M)$ which separates the graph. The dual to the conclusion of the theorem holds if and only if either the inside or outside of the ring consists of a single region. Thus, we have shown that Theorem 2.7 implies the following result of Birkhoff [**2**]:

2.9 THEOREM. *If M is an irreducible map, then M may* **not** *contain a ring of five regions unless they surround a pentagon.*

3. Reducible configurations. Theorem 2.9 of the last section suggests a definition:

3.1 DEFINITION: Let G be a graph. Then we call G a **reducible configuration** if G cannot occur as a subgraph of an irreducible graph. We define reducible configurations in maps using duality.

Thus, the previously mentioned result says that a ring of five regions not surrounding a pentagon is a reducible configuration.

We already have other types of reducible configurations; for example, any region with at most four sides. This allows us to derive a lower bound on the Birkhoff number.

If M is a cubic map and r_i denotes the number of regions bounded by i sides in the map, we have from Euler's formula

$$2m = \sum_i ir_i \,.$$

Putting these equations together yields the following well-known lemma (see, for example, Franklin [**3**, p. 154]):

3.2 LEMMA. *Let M be a cubic map. Then*

$$\sum_i (6 - i)r_i = 12.$$

If a map is irreducible, $r_i = 0$ for $i < 5$ and hence the only positive term in the

sum is $(6 - 5)\, r_5 = r_5$, the number of pentagons. We conclude immediately that any irreducible cubic map must have at least 12 pentagons.

If a map has exactly 12 pentagons, then it is a dodecahedron and can be 4-colored. Thus an irreducible map must have at least 13 regions. This proves that the Birkhoff number is at least 13.

To improve on this lower bound for the Birkhoff number, one must obtain more reducible configurations. Even then, however, increasing the lower bound can be very difficult because of the many combinatorial possibilities to be considered at every step.

Before listing other reducible configurations, we shall need some jargon. Our results will be in terms of vertices and degrees but of course can be dualized for regions and number of faces.

3.3 DEFINITION: We call a vertex v of degree k a k-**vertex** and write $d(v) = k$. Any vertex of degree 6 or less is called **minor**; vertices of degree 7 or more are called **major.** Let v_0 be a fixed vertex. A **neighbor** is a vertex adjacent to v_0. If a neighbor is a k-vertex, we call it a k-**neighbor.** Three vertices are in **triad** when they form the three corners of a triangle. Two neighbors of v_0 are **successive** when they form a triad with v_0. A vertex is **reducible** if it belongs to a reducible graph. A sequence $v_1, \cdots, v_r$ of neighbours of v_0 is called **successive** or **consecutive** if v_{i-1} and v_i are successive for $i = 1, \cdots, r$.

The following was one of the first reduction theorems:

3.4 THEOREM (Birkhoff [2]). *A 5-vertex is reducible when it has three consecutive 5-neighbors.*

Franklin [1] proved an analogous theorem about 6-vertices.

3.5 THEOREM. *A 6-vertex is reducible if it has three consecutive 5-neighbors.*

These results yield a corollary (Franklin [1]):

3.6 COROLLARY. *A 5-vertex v_0 is reducible when it has three 5-neighbors and a 6-neighbor.*

Proof: By Theorem 3.4, v_0 must have three consecutive neighbors v_1, v_2, v_3, where v_2 is a 6-vertex and v_1 and v_3 are 5-vertices or else v_0 is reducible. But now v_2 has three consecutive 5-neighbors v_1, v_0, v_3 so it is reducible by Theorem 3.5.

Franklin [2] also proved the following result:

3.7 THEOREM. *A 5-vertex with two 5-neighbors and three 6-neighbors is reducible.*

Winn [1] proved still another reduction theorem:

3.8 THEOREM. *A 5-vertex is reducible if it has one 5-neighbor and four 6-neighbors.*

Choinacki [1] and Winn [1] obtained another reduction result for 5-vertices.

3.9 THEOREM. *A 5-vertex all of whose neighbors are 6-vertices is reducible.*

Putting together the preceding results, we obtain the following corollary due to Winn [**1**]:

3.10 COROLLARY. *A 5-vertex is reducible when all of its neighbors are minor vertices.*

Thus, in an irreducible graph every 5-vertex is adjacent to a major vertex.

Bernhart ([**1**] and [**2**]) proved the following reduction theorem for a 6-vertex:

3.11 THEOREM. *A 6-vertex is reducible if it has three successive neighbors with degrees 5, 6, and 5, respectively.*

Winn [**1**] went on from there to obtain an analogue to Corollary 3.10 for 6-vertices.

3.12 THEOREM. *A 6-vertex is reducible when all of its neighbors are minor.*

Errera [**1**] obtained some general results about the number of consecutive 5-neighbors of an n-vertex in an irreducible graph.

3.13 THEOREM. *An n-vertex in an irreducible graph can have at most $n - 3$ consecutive 5-neighbors for n even and at most $n - 2$ for n odd.*

For $n = 7$, his result was improved by Winn [**2**].

3.14 THEOREM. *A 7-vertex with more than four consecutive 5-neighbors is reducible.*

Thus, a 7-vertex with six or more 5-neighbors is reducible; that is, in an irreducible graph, there are at most five 5-vertices adjacent to any 7-vertex.

Several new reducible configurations were discovered by Ore and Stemple [**1**]. For example, we have the following result:

3.15 THEOREM. *Let v_0 be a 5-vertex with neighbors v_1, v_2, v_3, v_4, v_5. If the corresponding list of degrees is (6, 5, 5, 6, 7) and v_4 and v_5 are in triad with a 5-vertex $w \neq v_0$, then the configuration is reducible.*

We have not attempted here to list all, or even nearly all, reducible configurations, but rather to give the flavor of the sorts of manipulations involved in obtaining them.

For a listing of most reducible configurations, see the paper of Ore and Stemple [**1**]. One may also consult Ore [**1**, Chapter 12] and Franklin [**3**, p. 156].

CHAPTER 4. RESULTS

1. Some sufficiency theorems. Any of the following conditions is sufficient to insure that a planar map be 4-colorable:

1.1 CONDITION: *Some region is bounded by at most 4 edges* (see Chapter 2, Section 1).

1.2 CONDITION: *Each region is bounded by at most five edges* (Aarts and de Groot [1]).

1.3 CONDITION: *There are at most* 21 *vertices of degree* 3 (Finck and Sachs [1]).

1.4 CONDITION: *There is at most one region of more than six sides and the map is irreducible* (Winn [1]).

1.5 CONDITION: *The countries with more than four neighbors can be divided into two classes such that one class has at most one country and no two countries in the other class are neighbors* (Dirac [8]).

1.6 CONDITION: *The number of edges in the boundary of each region is a multiple of* 3, *and the map is bridgeless cubic* (Winn [1]).

Very few constructions have been given which show how to color some general class of maps. The following scheme shows us how to 3-color the edges of a particular kind of map.

Let M be a cubic bridgeless map. Suppose that the number of edges in the boundary of every region is a multiple of 3. Ringel [1, p. 19] has given a constructive scheme for 3-coloring the edges of M.

Call the three colors 1, 2, and 3, and give them the usual cyclic ordering so that 2 follows 1, 3 follows 2, and 1 follows 3. If e, f, and g are the three edges of M incident with some vertex, give them the cyclic ordering induced by the clockwise orientation of the plane; that is, f follows e if, moving clockwise from e, we first encounter f.

1.7 COLORING SCHEME: Begin with some edge e of M and color it arbitrarily, say with 1. Now consider the four edges adjacent to e, two at each endpoint. In the cyclic orderings at each endpoint, these four edges either follow or precede e. Give them the corresponding color. (Thus, if f follows e, color f with 2.) Continue the process until all edges have been colored.

This procedure is unambiguous—in other words, only one color is assigned to each edge. Hence, no two adjacent edges receive the same color.

This provides us with a constructive proof of the sufficiency of Condition 1.6 since, given a 3-coloring of the edges of a cubic bridgeless map, we can then construct a 4-coloring of the regions of the map.

1.8 CONJECTURE C_{31}: *If a critical* 5-*chromatic graph contains a complete graph on three vertices, then the graph can be contracted to a complete graph on five vertices.*

The truth of this conjecture implies the truth of Conjecture C_1 (Dirac [5]). Conjecture C_1 implies Conjecture C_{23} of which this Conjecture is a special case.

1.9 THEOREM. *If* k (>2) *is the maximum degree of any vertex in a graph without loops and without complete subgraphs on* $k+1$ *vertices, then the graph is* k-*colorable.*

This is the famous result of Brooks [1] which contains the dual of Condition 1.2

as a corollary. The following results indicate that k-chromatic graphs may be somewhat pathological.

1.10 THEOREM. *For any $k > 1$ there exists a k-chromatic graph which has no circuit (region) of less than* 6 *edges* (B. Descartes [**1**]).

1.11 THEOREM. *If $d \geqq k \geqq 2$, then there exist regular connected k-chromatic graphs of degree d and of an arbitrarily large number of vertices* (Dirac [**4**]).

For $k \geqq 4$ Dirac constructs a k-chromatic graph which does not contain a complete k-graph as a subgraph and in which the degree of every vertex except one is $k - 1$.

2. Coloring problems on surfaces other than the plane. In view of the fact that the four-color problem is unsolved, it is perhaps surprising that the analogous problems on other orientable surfaces have been solved completely!

2.1 DEFINITION: A surface is said to have **genus** p if it is a homeomorph of a sphere with p handles.

2.2 THEOREM. *For any positive integer p, the chromatic number of a graph embedded in the (orientable) surface of genus p is at most χ_p where*

$$\chi_p = \left[\frac{7 + \sqrt{1 + 48p}}{2}\right].$$

This is Heawood's Map-Coloring Theorem—see Busacker and Saaty [**1**, p. 94] for the proof. Note that if this theorem held for $p = 0$, we would have a proof of Conjecture C_1. Unfortunately, the only known proof of Theorem 2.2 depends on having $p > 0$.

Recently, Ringel and Youngs [**2**] have shown that if $p \geqq 1$, then there always exists a graph which can be embedded in the surface of genus p whose chromatic number is exactly equal to χ_p (see also Youngs [**1**] and Berge [**1**, p. 218]).

We might also mention here that Ringel [**2**] has given an interesting six-color problem on the sphere in which he asks for a coloring of both regions and vertices using 6 colors so that no 2 adjacent vertices or regions are colored the same and so that no vertex receives the same color as the regions on whose boundaries it lies.

3. One, two, and three and more colorability. Clearly a graph is 1-colorable if and only if it consists of isolated vertices (i.e., it is totally disconnected).

3.1 THEOREM. *A map is properly colorable with two colors if and only if every vertex is of even degree.*

This follows from the fact that a graph is bipartite if and only if it has no circuits of odd length (König [**1**, p. 151]).

3.2 THEOREM. *A cubic map is properly colorable with three colors if and only if each region is bounded by an even number of edges* (Franklin [**3**, p. 198]).

Dually, a maximal planar graph is 3-colorable (i.e., 3-partite) if and only if every vertex has even degree. Unfortunately, no general useful characterization of 3-partite graphs or 3-partite planar graphs is known at present.

3.3 THEOREM. *The edges of a cubic map can be properly colored with four colors* (Golovina and Yaglom [**1**, p. 43]).

This is also a corollary of the Shannon-Vizing bound on the chromatic index.

Grünbaum [**1**] has shown that every planar map with less than 4 triangles is 3-colorable. As a consequence of the theorem of Brooks, triangular maps (other than the tetrahedron) are 3-colorable.

3.4 THEOREM. *If a triangular map can be properly colored with two colors, then its vertices can be properly colored with three colors.*

See Dynkin and Uspenskii [**1**].

3.5 THEOREM. *The edges of a cubic map can be colored with two colors α and β so that each vertex is incident with one edge colored with α and two edges colored with β.*

This theorem is due to Petersen [**1**]. It can be restated in the form: Every bridgeless cubic map is the sum of a 1-factor and a 2-factor. Petersen gave an example to show that a similar result with three 1-factors cannot be obtained. (See Fig. 7.)

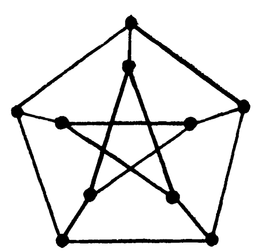


FIG. 7

Marathe [**1**] has shown that Petersen's theorem is a corollary of the following result:

3.6 THEOREM. *Any triangular map with an even number of triangles can be colored with two colors α and β so that each triangle is bounded by one edge colored α and two edges colored β.*

4. The sufficiency of six colors. We already know that 5 colors suffice to color any planar map, but we shall give a short direct proof here that 6 colors suffice since the argument demonstrates, once again, the ubiquity of Euler's formula in these coloring problems and since it gives us a method for reducing the number of regions in a cubic map.

Consider Euler's formula $n - m + r = 2$ and substitute $n = 2m/3$ (for a cubic map). This gives $6(r - 2) = 2m$. Since $6r > 6(r - 2) = 2m$ we prove that 6 colors are sufficient to color any cubic map. This is clear when $r < 6$. If $r \geqq 6$, then there must be (as we already know) at least one region bounded by 5 or less edges. Applying induction, we may assume that all maps are 6-colorable for $r - 1$ regions. If we remove a less than six sided region of the map and extend the edges of its neighbors in such a way that each vertex is of degree three and the entire removed region is covered by its five neighbors as in the diagram below, we can 6-color the map and then reinstate the removed region, coloring it with the sixth color not appearing in any of its five neighbors. (See Fig. 8.)

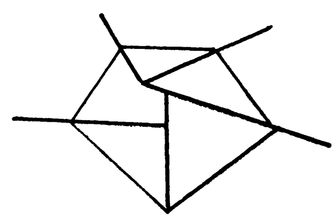


FIG. 8

5. The uniqueness of colorings. The uniqueness of the colorability of a graph has also been investigated. A complete presentation is given in the book by Harary [**2**, p. 137]. Note that in a unique coloring, each vertex must be adjacent to vertices whose totality is colored with all the remaining colors (at least once). We have the following results for uniqueness of coloring with k colors:

5.1 THEOREM. *In the partition of the vertices into subsets induced by the coloring, the vertices of every pair of subsets with their connecting edges form a connected subgraph* (Cartwright and Harary [**1**]).

5.2 THEOREM. *The graph is* $(k - 1)$*-connected. The corresponding subgraph for* m *subsets,* $2 \leqq m \leqq k$ *is* $(m - 1)$*-connected.*

5.3 THEOREM. *For each* $k \geqq 3$ *there is a uniquely* k*-colorable graph with no subgraph isomorphic to the complete graph on* k *vertices* (Harary, Hedetniemi, and Robinson [**1**]).

It is also known (Chartrand and Geller [**1**]) that no planar graph is uniquely 5-colorable; every uniquely 4-colorable planar graph is maximal planar; and that a planar 3-colorable graph in which each vertex belongs to the last triangle of a linear sequence of triangles each sharing an edge with its immediate neighbors is uniquely 3-colorable. A uniquely 3-colorable planar graph on $n \geqq 4$ vertices contains at least two triangles.

In general, the coloring of a map or a graph is not unique. There are a number

of papers studying the number of colored graphs. We give a sample of the known ones in addition to the discussion of chromatic polynomials already given.

Let $F_n(k)$ denote the total number of k-colored graphs on n labelled vertices and let $M_n(k)$ denote the number of graphs on n vertices that are colored in at most k colors; also let $f_n(k)$ denote the number of connected k-colored graphs on n vertices. Read [1] gives:

$$\sum_{n=1}^{\infty} 2^{-\frac{1}{2}n^2} F_n(k) \frac{x^n}{n!} = \left\{ \sum_{s=1}^{\infty} 2^{-\frac{1}{2}s^2} \frac{x^s}{s!} \right\}^k,$$

$$\sum_{n=0}^{\infty} 2^{-\frac{1}{2}n^2} M_n(k) \frac{x^n}{n!} = \left\{ \sum_{s=0}^{\infty} 2^{-\frac{1}{2}s^2} \frac{x^s}{s!} \right\}^k,$$

$$1 + \sum_{n=1}^{\infty} 2^{-\frac{1}{2}n^2} F_n(k) \frac{x^n}{n!} = \left\{ \sum_{n=1}^{\infty} f_n(k) \frac{x^n}{n!} \right\}.$$

Wright [1] has proved some asymptotic formulas for $F_n(k)$, $M_n(k)$, $f_n(k)$. Carlitz [1] has analyzed some arithmetic properties of these numbers. An interesting and rather simple one to quote is:

$$M_n(k) \equiv k \,(\text{mod } 2^n) \qquad (n > 2)$$

from which it follows that $M_n(k)$ is odd if and only if k is odd.

5.4 DEFINITION: A map is **rooted** when a vertex, an edge and a face that are mutually incident are specified as **root-vertex, root-edge** and **root-face**, respectively.

Consider a bridgeless cubic rooted map with $2n$ vertices. Two colorings are not considered as distinct if they differ only by a permutation of the four colors. Suppose that the root-face is red, the other face incident with the root-edge is blue, the third face incident with the root-vertex green, and the fourth color, yellow.

5.5 DEFINITION: The Tait cycle separating blue and green from red and yellow is called the **basic Tait cycle** of the coloring (it passes through the root-edge).

5.6 DEFINITION: The **rank** of the coloring is equal to the number of components of the basic Tait cycle minus one.

W. T. Tutte [7, 9] has shown that the average number of 4-colorings for such maps with 2 n-vertices is asymptotically equal to the following expressions:

$$8(3\pi n)^{-\frac{1}{2}}(32/27)^n \qquad \text{for rank 0,}$$

$$8(3\pi n)^{-\frac{1}{2}}(4/\pi - 1)n^{\frac{1}{2}}(32/27)^n \text{ for rank 1.}$$

One can also introduce the notion of semi -uniquely 4-colorable graphs.

5.7 DEFINITION: Suppose $\chi(G) = 4$. Let v and w be vertices of G. Then we say that v and w are **brothers** if any 4-coloring of G assigns the same colors to v and w. We say that G is **semi-uniquely 4-colorable** if it has a pair of vertices which are brothers.

D. L. Greenwell [1] has proved that the following conjecture is equivalent to Conjecture C_1:

5.8 CONJECTURE C_{32}: *Let G be a semi-uniquely 4-colorable planar graph and let v and w be a pair of brothers in G. Then the graph G′ obtained from G by joining v and w with an edge is not planar.*

6. Some recent developments. It would be totally beyond the scope of this paper to discuss the problem of coloring infinite planar graphs. We might mention here, however, some recent work of R. Halin [3] on coloring numbers which has applications to finite graphs. The **coloring number,** col (G), of a (possibly infinite) graph G was first introduced by Erdös and Hajnal [1] and is defined as the smallest cardinal k for which there exists a well-ordering of the vertices of G such that every vertex v of G is adjacent to less than k vertices preceding it in the ordering. Clearly, $\chi(G) \leqq \text{col}(G)$. Halin shows that if col (G) is sufficiently large, then G must contain subdivisions of any complete graph on fewer than col (G) vertices.

We should also like to draw the reader's attention to some other recent papers. S. Hedetniemi [1] defines a **disconnected-coloring** (or **D-coloring**) of a graph $G = (V, E)$ as a partition $V = V_1 \cup \cdots \cup V_n$ of V such that, for every i, the section graph of G induced by the subset V_i is disconnected. The **D-chromatic** number $\chi_d(G)$ is the smallest number of subsets in any D-coloring of G. The D-chromatic number shares many properties with the chromatic number but differs in others. For example, Hedetniemi gives the following theorem:

6.1 THEOREM. *If G is planar, then* $\chi_d(G) \leqq 4$.

Other recent results have dealt with edge coloring. M. Rosenfeld [1] proved the following theorem:

6.2 THEOREM. *Let G be a cubic graph with n vertices. Then G is homomorphic to a Tait-colorable cubic graph G′ with* $(6n + 5)/5$ *vertices.*

In a recent paper, M. R. Williams [1] suggests an improvement of a heuristic coloring procedure developed by Peck and Williams [1]. The latter procedure takes a graph and proceeds as follows to determine which vertices should be colored with the kth color (cf. Welsh and Powell [1]).

(i) Find the uncolored vertex v of highest degree.
(ii) Check to see if v is adjacent to any vertex already colored with the kth color.
(iii) If not, then color v with color k.
(iv) If yes, then remove v from consideration for color k and return to step (i).

This heuristic procedure uses the vector d whose ith component is the degree of the ith vertex. Williams modifies the above procedure by replacing d with a vector d^m defined recursively by setting $d = d^1$ and $d^{m+1} = Ad^m$, where A is the adjacency or vertex-vertex matrix of G. The vectors d^m converge to the dominant eigenvector

of A as $m \to \infty$. Williams observes that convergence generally occurs after $m = \sqrt[3]{n}$ iterations where n is the number of vertices in the graph.

Williams used his modified heuristic to color one graph of over 700 vertices using 28 colors. The graph was later found to contain a complete subgraph on 26 vertices so Williams' estimate was certainly not too high!

Striking out into other new directions, J. W. T. Youngs **[2]** indicates how his joint work with Ringel (Ringel and Youngs **[2]**), in which they settled the Heawood Conjecture, can be used to provide "slick" proofs that various conjectures, e.g., Conjecture C_4, are equivalent with the four color conjecture. Hopefully, these methods (current graphs, graphs with rotation, Kirchhoff's Law) will eventually provide us with some new information in this area although they have not yet done so.

Finally, we should like to mention some recent work of ours with P. Kainen **[1]** in which we have considered the problem of relative colorings. Suppose we consider some planar graph G with a section subgraph, G', that has already been colored. A **relative coloring** of (G, G'), with respect to the given coloring of G', is a coloring of G which agrees with the given coloring on the vertices of G'.

Note that if G' is 4-colored, we may need as many as 4 new colors to color G relative to the coloring of G'. Let us write $\chi(G, G')$ for the maximum number of new colors needed in any relative coloring of (G, G'). We call this the **relative chromatic number** of (G, G').

We prove that the following conjecture is equivalent to the four color conjecture.

6.3 CONJECTURE C_{33}: *For any pair (G, G') with G planar and G' a (possibly empty) subgraph of G, we have $\chi(G, G') \leqq 4$.*

If we require G' to be connected, then we know of no examples where $\chi(G, G') > 3$. This leads us to make the following conjecture which implies Conjecture C_1.

6.4 CONJECTURE C_{34}: *For any pair (G, G') with G planar and G' a connected subgraph of G, we have $\chi(G, G') \leqq 3$.*

We do not know whether this conjecture is implied by the four-color conjecture.

Conclusion. To conclude, it may be of interest to give a quotation from a paper by a great living geometer, H. S. M. Coxeter **[1]**:

> If I may be so bold as to make a conjecture, I would guess that a map requiring five colors may be possible, but that the simplest such map has so many faces (maybe hundreds or thousands) that nobody, confronted with it, would have the patience to make all the necessary tests that would be required to exclude the possibility of coloring it with four colors. Many people believe, on the other hand, that the four-color theorem may be true; in fact, editors of journals often have the unhappy experience of receiving manuscripts in which it is "proved." Such manuscripts are either obviously incompetent or else so lengthy that the referee has a tedious job finding the flaw. The problem

has been considered by so many able mathematicians that anyone who can prove that a particular map really needs five, will become world-famous overnight.

There is still great and lively interest in the problem: Shimamoto of the Brookhaven National Laboratory Computer Center, is presenting a paper on a proof of the four-color problem. One of the steps in the proof depends on a complicated computer program which is still being worked on at this time.

My heartfelt thanks to my colleague and friend, Paul Kainen, for careful reading and suggestions which enriched the manuscript. I would also like to thank Michael Albertson and David Burman for help in obtaining information and Marilyn Dalick for her great patience in typing many versions of the manuscript over the past two years.

References

[1] J. M. Aarts and J. A. de Groot, A case of coloration in the four color problem, Nieuw Arch. Wisk., 11 (1963) 10–18.

[1] W. W. Rouse Ball and H. S. M. Coxeter, Mathematical Recreations and Essays, Macmillan, New York, 1947.

[1] C. Berge, The Theory of Graphs and its Application, Dunod, Paris, 1958 (in French); Methuen, London, 1962 (in English).

[1] A. Bernhart, Six-rings in minimal five-color maps, Amer. J. Math., 69 (1947) 391–412.

[2] ———, Another reducible edge configuration, Amer. J. Math., 70 (1948) 144–146.

[1] G. D. Birkhoff, A determinant formula for the number of ways of colouring a map, Ann. Math., 14 (1912) 42–46.

[2] ———, The reducibility of maps, Amer. J. Math., 35 (1913) 115.

[3] ———, On the number of ways of coloring a map, Proc. Edinburgh Math. Soc., 2 (1930) 83–91.

[4] ———, On the polynomial expressions for the number of ways of coloring a map, Ann. Scuola Norm. Sup. Pisa, 2 (1934) 85–103.

[1] G. D. Birkhoff and D. Lewis, Chromatic polynomials, Trans. Amer. Math. Soc., 60 (1946) 355–451.

[1] R. L. Brooks, On colouring the nodes of a network, Proc. Cambridge Philos. Soc., 37 (1941) 194–197.

[1] R. Busacker and T. L. Saaty, Finite Graphs and Networks: An Introduction with Applications, McGraw-Hill, New York, 1965.

[1] L. Carlitz, The number of colored graphs, Canad. J. Math., 15 (1963) 304–312.

[1] D. Cartwright and F. Harary, The number of lines in a digraph of each connectedness category, SIAM Review, 3 (1961) 309–314.

[2] ——— , On the coloring of signed graphs, Elem. Math., 23 (1968) 85–89.

[1] A. Cayley, On the colouring of maps, Proc. London Math. Soc., 9 (1878) 148. See also Proc. Roy. Geographical Soc., 1 (1879) 259.

[1] G. Chartrand and D. Geller, Uniquely colorable planar graphs, J. Combinatorial Theory, 6 (1969) 271–278.

[1] C. A. Choinacki, A contribution to the four color problem, Amer. J. Math., 64 (1942) 36–54.

[1] J. Chuard, Les réseaux cubiques et le problème des quatre couleurs, Mém. Soc. Vaudoise Sci. Nat., No. 25, 4 (1932) 41–101.

[1] H. Coxeter, The four-color map problem, Math. Teacher, 52 (1959) 283–289.

[1] George B. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, N. J. 1963.

[1] N. Debruijn, A color theorem for infinite graphs and a problem in the theory of relations, Nederl. Akad. Wetensch. Proc. Ser. A 54 Indag. Math., 13 (1951) 371.

[1] B. Descartes, Solution to problem 4526, this MONTHLY, 61 (1954) 352.

[1] B. Descartes and R. Descartes, La coloration des Cartes, Eureka, 31 (1968) 29–31.

[1] G. A. Dirac, Note on the colouring of graphs, Math. Zeitschr., 54 (1951) 347–353.

[2] ———, A property of 4-chromatic graphs and some remarks on critical graphs, J. London Math. Soc., 27 (1952) 85-92.

[3] ———, Some theorems on abstract graphs, Proc. London Math. Soc., Ser. 3, 2, (1952) 69–81.

[4] ———, The structure of k-chromatic graphs, Fund. Math., 40 (1953) 42–55.

[5] ———, Theorems related to the four colour conjecture, J. London Math. Soc., 29 (1954) 143–149.

[6] ———, Circuits in critical graphs, Monatsh., Math., 59 (1955) 178–187.

[7] ———, Map colour theorems related to the Heawood colour formula, J. London Math. Soc., 31 (1956) 460–471.

[8] ———, A theorem of R. L. Brooks and a conjecture of H. Hadwiger, Proc. London Math. Soc., 7 (1957) 161–195.

[9] ———, Trennende Knotenpunktmengen und Reduzibilität abstrakter Graphen mit Anwendung auf das Vierfarbenproblem, J. für Math., 204 (1960) 116–131.

[10] ———, On the structure of 5 and 6-chromatic abstract graphs, J. für Math., 214 (1964) 43–52.

[1] E. B. Dynkin and W. A. Uspenskii, Multicolor Problems, Heath, Boston, 1963; In German, Berlin, 1955; 1st Russian edition, 1952.

[1] A. Errera, Une contribution au problème des quatre couleurs, Bull. de la Soc. Math. de France, 53 (1925) 42.

[1] P. Erdös and A. Hajnal, On chromatic numbers of graphs and set-systems, Acta Math. Acad. Sci. Hungar, 171 (1966) 61–99.

[1] A. P. Ershov and G. I. Kozhukhin, Estimates of the chromatic number of connected graphs, Dokl. Akad. Nauk, 142 (1962) 270–273; Trans. Soviet Math., 3 (1962) 50–53.

[1] H. J. Finck and H. Sachs, Über eine von H. S. Wilf angegebene Schranke für die chromatische Zahl endlicher Graphen, Math. Nachr., 39 (1969) 373–386.

[1] P. Franklin, The four color problem, Amer. J. Math., 44 (1922) 225–236.

[2] ———, Note on the four color theorem, J. Math. Phys., 16 (1938) 172–184.

[3] ———, The four color problem, Scripta Math., 6 (1939) 149–156 and 197–210.

[1] T. Gallai, Kritische Graphen, I and II, Publ. Math. Inst. Hungarian Acad. Sci. A., 8 (1963) 165–192; 9 (1964) 373–395.

[1] L. I. Golovina and I. M. Yaglom, Induction in Geometry, Heath, Boston, 1963.

[1] D. L. Greenwell, Semi-uniquely n-colorable graphs, Proc. Second Louisiana Conference in Combinatorics and Graph Theory, to appear.

[1] B. Grünbaum, Grötzsch's theorem on 3-colorings, Michigan Math. J., 10 (1963) 303–310.

[2] ———, Convex Polytopes, Interscience Publishers, New York, 1967.

[1] H. Hadwiger, Über eine Klassifikation der Streckenkomplexe, Vierteljschr. Naturforsch. Ges., Zürich, 88 (1943) 133–142.

[2] ———, Ungelöste Probleme, Elem. Math., 12 (1957) 61–62.

[3] ———, Ungelöste Probleme, Elem. Math., 13 (1958) 127–128.

[1] G. Hajös, Über eine Konstruktion nicht n-färberer Graphen, Wiss. Zeitschr., Martin Luther Univ. Halle-Wittenburg, A 10 (1961) 116–117.

[1] R. Halin, Bemerkungen über ebene Graphen, Math. Annalen, 153 (1964) 38–46.

[2] ———, On a theorem of Wagner related to the four-color problem, homomorphisms, (German) Math. Ann., 153 (1964) 47–62.

[3] ———, A colour problem for infinite graphs, in Combinatorial Structures and Their Applications, Gordon and Breach, New York, 1970, 123.

[1] F. Harary, A complementary problem on non-planar graphs, Math. Mag., 35 (1962) 301–304.

[2] ———, Graph Theory, Addison-Wesley, Reading, Mass., 1969.

[1] F. Harary, S. T. Hedetniemi, and R. W. Robinson, Uniquely colorable graphs, J. Combinatorial Theory, 6 (April, 1969).

[1] F. Harary and W. T. Tutte, A dual form of Kuratowski's theorem, Canad. Math. Bull., 8 (1965), 17–20, 373.

[1] P. J. Heawood, Map-colour theorems, Quart. J. Math., Oxford Ser. 24 (1890) 332–338. See also G. A. Dirac, and Percy John Heawood, J. London Math. Soc., 38 (1963) 263–277.

[2] ———, On the four-colour map theorem, Quart. J. Math., 29 (1898) 270-285.

[3] ———, On extended congruences connected with the four-colour map theorem, Proc. London Math. Soc., 33 (1932) 252–286.

[4] ———, Failures in congruences connected with the four-colour map theorem, Proc. London Math. Soc., 40 (1936) 189–202.

[5] ———, Note on a correction in a paper on map-congruences, J. London Math. Soc., 18 (1943) 160–167; 19 (1944) 18–22.

[1] S. Hedetniemi, Disconnected-colorings of graphs, in Combinatorial Structure and their Application, Gordon and Breach, New York, 1970, 163.

[1] P. C. Kainen and T. L. Saaty, Relative colorings of graphs, to appear.

[1] J. B. Kelly and L. M. Kelly, Paths and circuits in critical graphs, Amer. J. Math., 76 (1954) 786–792.

[1] A. B. Kempe, On the geographical problem of the four colors, Amer. J. Math., 2 (1879) 193–200.

[1] D. König, Theorie der endlichen und unendlichen Graphen, Leipzig, 1936; reprinted Chelsea, New York, 1950.

[1] K. Kuratowski, Sur le problème des courbes gauches en topologie, Fund. Math., 15 (1930) 271–283.

[1] C. L. Liu, Introduction to Combinatorial Mathematics, McGraw-Hill, New York, 1968.

[1] C. R. Marathe, On the dual of a trivalent map, this MONTHLY, 68 (1961) 448–455.

[1] K. O. May, The origin of the four-color conjecture, Isis, 56 (1965) 346–348.

[1] G. J. Minty, A theorem on n-coloring the points of a linear graph, this MONTHLY, 69 (1962) 623–624.

[1] E. Nordhaus and J. Gaddum, On complementary graphs, this MONTHLY, 63 (1956) 175–177.

[1] O. Ore, The Four Color Problem, Academic Press, New York, 1967.

[2] ———, The Theory of Graphs, A. M. S. Colloquium Publications, 1962.

[1] O. Ore and G. J. Stemple, Numerical calculations on the four-color problem, J. Combinatorial Theory, 8 (1970) 65–78.

[1] Erika Pannwitz, Review of Chuard [1] in Jahrbuch über die Fortschritte der Math., 58 (1932) 1204.

[1] J. E. L. Peck and M. R. Williams, Examination scheduling, Algorithm 286, Comm. ACM, 9 (6) (1966).

[1] J. Petersen, Die Theorie der regulären Graphen, Acta Math., Stockholm, 15 (1891) 193–220. See also Intermed. Math., 5 (1898) 225–227; *ibid.*, 6 (1899) 36–38.

[1] L. Posa, A theorem concerning Hamiltonian lines, Publ. Math. Inst. Hungar. Acad. Sci., 7 (1962) 225–226.

[1] R. C. Read, The number of k-coloured graphs, Canad. J. Math., 12 (1960) 410–414.

[2] ———, An introduction to chromatic polynomials, J. Combinatorial Theory, 4 (1968) 52–71.

[1] C. N. Reynolds, On the problem of coloring maps in four colors, 1, Ann. Math., 28 (1926–27) 477–492.

[1] G. Ringel, Färbungsprobleme auf Flächen und Graphen, Berlin, 1959.

[2] ———, A six-color problem on the sphere, (German) Abh. Math. Sem. Univ. Hamburg, 29 (1965) 107–117.

[1] G. Ringel and J. W. T. Youngs, Remarks on the Heawood conjecture, Proof Techniques in Graph Theory (F. Harary, ed.), Academic Press, New York, 1969.

[2] ——————, Solution of the Heawood map coloring problem, Proc. Nat'l Acad. Sci., 60 (1968) 438–445.

[1] A. van Rooij and H. S. Wilf, The interchange graphs of a finite graph, Acta Math, Sci. Hungar., 16 (1965) 263–269.

[1] M. Rosenfeld, On Tait coloring of cubic graphs, in Combinatorial Structures and Their Applications, Gordon and Breach, New York, 1970, 373.

[1] G. C. Rota, On the foundations of combinatorial theory; theory of Möbius functions, 2, Wahrscheinlichkeitstheorie und Verw. Gebiete, 2 (1964) 240–368.

[1] B. Roy, Algèbre Moderne et Théorie des Graphes, Dunod, Paris, two volumes, 1969 and 1970.

[1] T. L. Saaty, Remarks on the four color problem, the Kempe catastrophe, Math. Mag., 40 (January 1967) 31–36.

[1] J. Sédlaček, Some properties of interchange graphs, Theory of Graphs and Its Applications, Symposium Smolenice, 1963, 145–150.

[1] C. E. Shannon, A theorem on coloring the lines of a network, J. Math. and Phys., 28 (1949).

[1] E. Steinitz and H. Rademacher, Vorlesungen über die Theorie der Polyeder, Berlin, 1934.

[1] W. E. Story, Note on Mr. Kempe's paper on the geographical problem of the four colours, Amer. J. Math., 2 (1879) 201–204.

[1] G. Szekeres and H. S. Wilf, An inequality for the chromatic number of a graph, J. Combinatorial Theory, 4 (1968) 1–3.

[1] P. G. Tait, Remarks on the colouring of maps, Proc. Roy. Soc., Edinburgh, 10 (1880) 729.

[2] ———, Note on a theorem in geometry of position, Trans. Roy. Soc., Edinburgh, 29 (1880) 657–660.

[3] ———, On Listing's "Topologie", Phil. Mag., 17 (1884) 30–46.

[1] J. M. Thomas, The Four Color Theorem, 60 Slocum Street, Philadelphia, Pennsylvania, 1969.

[1] W. T. Tutte, On Hamiltonian circuits, J. London Math. Soc., 21 (1946) 98–101.

[2] ———, The factors of graphs, Canad. J. Math., 4 (1952) 314.

[3] ———, A theorem on planar graphs, Trans Amer. Math. Soc., 82 (1956) 99–116.

[4] ———, A non-Hamiltonian graph, Canad. Math. Bull., 3 (1960) 1–5.

[5] ———, On the algebraic theory of graph colorings, J. Combinatorial Theory, 1 (June, 1966).

[6] ———, A geometrical version of the four color problem, Proc. of the Conference held at University of North Carolina, Chapel Hill, April 10–14, 1967.

[7] ———, On the enumeration of four-colored maps, SIAM, J. Appl. Math., (March 1969) 454–460.

[8] ———, On chromatic polynomials and the golden ratio, J. Combinatorial Theory, 9 (October, 1970).

[9] ———, Even and odd 4-colorings, Proof Techniques in Graph Theory, Academic Press, New York, 1969.

[10] ———, The golden ratio in the theory of chromatic polynomials, Annals of the New York Academy of Sciences, 175 (1970), 391–402.

[11] ———, The Connectivity of Graphs, Toronto University Press, Toronto, 1967.

[12] ———, More about chromatic polynomials and the golden ratio, Combinatorial Structures and Their Applications, Gordon and Breach, New York, 1969, 439.

[1] O. Veblen, An application of modular equations in analysis situs, Ann. Math., 14 (1912–1913) 86–94.

[2] ———, Analysis situs, American Math. Soc., 5 Cambridge, (1922); 2nd edition, New York, (1931).

[1] V. G. Vizing, On an estimate of the chromatic class of a p-graph, (Russian) Diskret. Analiz., 3 (1964) 25–30.

[2] ———, Chromatic class of multigraph, Théorie des Graphes, Journée Internationale d'Etude, Rome, 3 (1966) 29–33.

[3] ———, On the number of edges in a graph with given radius, (Russian) Dokl. Akad. Nauk. SSSR, 173 (1967) 1245–1246.

[1] K. Wagner, Ein Satz über Komplexe, J. -ber. Deutsch. Math-Verein, 46 (1936) 21–22.

[2] ———, Bemerkungen zum Vierfarbenproblem, J. -ber. Deut. Math.-Ver., 46 (1936) 26–32.

[3] ———, Über eine Eigenschaft der ebenen Komplexe, Math. Ann., 114 (1937) 570–590.

[4] ———, Beweis einer Abschwächung der Hadwiger-Vermutung, Math. Ann., 153 (1964) 139–141.

[1] M. E. Watkins, A theorem on Tait colorings with an application to the generalized Petersen graphs, J. Combinatorial Theory, 6 (1969).

[1] D. J. Welsh and M. B. Powell, An upper bound for the chromatic number of a graph and its application to time-tabling problems, Computer Journal, 10 (1967) 85–86.

[1] H. Whitney, A theorem on graphs, Ann. Math., 32 (1931) 378–390.

[2] ———, The coloring of graphs, Ann. Math., 33 (1932) 688–718.

[3] ———, Non-separable and planar graphs, Trans. Amer. Math. Soc., 34 (1932) 339–362.

[4] ———, A logical expansion in mathematics, Bull. Amer. Math. Soc., 38 (1932) 572–579.

[5] ———, Planar graphs, Fund. Math., 21 (1933) 73–84.

[6] ———, Isomorphic graphs, Amer. J. Math., 55 (1933) 245–254.

[7] ———, A numerical equivalent of the four color map problem, Monatsh. Math. und Physic, (1937) 207–213.

[8] ———, Congruent graphs and the connectivity of graphs, Amer. J. Math., 54 (1932) 150–168.

[1] H. S. Wilf, The eigenvalues of a graph and its chromatic number, J. London Math. Soc., 42 (1967) 330–332.

[2] ———, Hadamard determinants, Möbius functions and the chromatic number of a graph, Bull. Amer. Math. Soc., (September, 1968) 960–964.

[1] M. R. Williams, A graph theory model for the computer solution of university time-tables and related problems, Ph. D. thesis, University of Glasgow, 1969.

[1] C. E. Winn, A case of coloration in the four color problem, Amer. J. Math., 59 (1937) 515–528.

[2] ———, On certain reductions in the four color problem, J. Mathematical Phys., 16 (1938) 159–171.

[3] ———, Sur l'historique du problème des quatre couleurs, Bull. Inst. Egypte, 20 (1939) 191–192.

[4] ———, On the minimum number of polygons in an irreducible map, Amer. J. Math. 62 (1940) 406–416.

[1] E. Wright, Counting colored graphs, Canad. J. Math., 13 (1961) 683–693.

[1] H. Yamabe and D. Pope, A computational approach to the four-color problem, Math. Comp., (1961) 250–253.

[1] J. W. T. Youngs, The Heawood map colouring conjecture, in Graph Theory and Theoretical Physics, (F. Harary, ed.), Academic Press, London, 1967, 313–354.

[2] ———, Remarks on the four color problem, Combinatorial Structures and Their Applications, Gordon and Breach, New York, 1970, 479.

[1] E. C. Zeeman, Seminar on Piecewise-Linear Topology (mimeographed), Inst. des Hautes Études Scientifiques, 1965.

EXPLICIT FORMULAS FOR BERNOULLI NUMBERS

H. W. GOULD, West Virginia University

A recent paper by Higgins [19] offers what is purported to be a new finite double series for the Bernoulli numbers with similar results for the Euler numbers. The paper gives an introductory account of the history of the Bernoulli numbers and quotes from some very old and authoritative sources, as well as recent papers about the numerical computation of the Bernoulli, Euler, and Tangent numbers. However the author seems to have missed other equally valuable papers so that he is constrained to state that "as far as I am aware there has been no explicit evaluation of them apart from this [an integral given by Whittaker and Watson], though values have from time to time been tabulated." The object of the present paper is to set matters straight by presenting a bibliography on explicit formulas for the Bernoulli numbers, and show how one can easily manufacture expressions for these numbers.

Basically, what Higgins found when $a = 0$ in his general formula (2.5) is

$$B_n = \sum_{k=0}^{n} \frac{1}{k+1} \sum_{j=0}^{k} (-1)^j \binom{k}{j} j^n, \qquad n \geqq 0, \tag{1}$$

and the reader will have no difficulty in seeing that the lower limits of summation in both cases may be replaced by $k = 1$ and $j = 1$ so as to agree with the form in which Higgins gives the result, a form valid for $n \geqq 1$. The formula is quite old, and it is difficult to say how to assign priorities, but the interested reader should consult the sources listed here with special attention to the book by Saalschütz [24]. Saalschütz gives [pp. 54–116] a total of 38 explicit formulas for the Bernoulli numbers, usually giving some reference in the older literature together with a proof. The notations used are quite different from recent ones, and the dozens of different notations in use for the numbers of Bernoulli and Stirling, etc., is surely one explanation for the formulas not being widely known. Yet each notation has its own elegance and place.

The book of Saalschütz has been out of print for many decades and has become quite hard to locate, with only a very exceptional library having a copy. The present writer was able to persuade University Microfilms to track down a copy and now a Xerographed version can be gotten from them very easily. The problem was two-fold: question of possible copyright and availability of a copy to photograph. Both problems were solved. The copy from Yale University Library was used. Anyone wishing to work with Bernoulli numbers in any great detail ought to examine the book.

In a recent book review [17] I called attention to formula (1) and deplored the fact that almost no current books on infinite series ever mention or derive (1), and even Knopp in his famous booklet on series asserted that the Bernoulli numbers "cannot be specified by means of a simple formula — except, say, by means of a determinant..." such is the widespread misinformation at hand.

Higgins cites a paper by von Staudt but misses one published five years later [28] in which explicit formulas are given. The formula (1) may be found derived in the

well-known book by Jordan [20] on finite differences... see p. 236 there. The uses of (1) in connection with the von Staudt-Clausen theorem may be read in the papers of Carlitz cited here. Particular notice is called to Carlitz's paper [8] where the formula (1) is derived in two ways, by infinite series and by finite differences, and then applied to arithmetic studies.

Garabedian [13] rediscovered the formula

$$B_{n+1} = \frac{(-1)^{n+1}(n+1)}{2^{n+1}-1} \sum_{j=0}^{n} (-1)^j 2^{-j-1} \Delta^j 1^n, \tag{2}$$

which Carlitz remarked [6] as being a very old result. In fact he traced it back to the paper by Worpitzky [31]. Carlitz then showed how one could easily obtain an equally nice formula giving B_n as a linear combination of differences of zero.

Shanks [27] also rediscovered a common formula for the Bernoulli numbers, and a discussion of the related numbers of Euler and Worpitzky (and Nielsen) was given by Carlitz [5].

Formula (1) was posed as a problem by Burger [2] who used an expansion of $\exp(t(e^z - 1))$ to obtain the result.

Munch [22] found an old result and published the formula in the form

$$B_n = \frac{1}{n+1} \sum_{k=1}^{n} \sum_{j=1}^{k} (-1)^j j^n \frac{\binom{n+1}{k-j}}{\binom{n}{k}}. \tag{3}$$

Formulas similar to (2) may be found in Schwatt's book [25]. Schwatt's book is a valuable source of technique and formulas, yet little known. Its importance was called to the attention of the present author by Carlitz some years ago, and because of the limited availability of the book the present writer was delighted to be able to persuade the Chelsea Publishing Company to reprint the book, using some corrections noted by this writer. However, as with the book of Nielsen cited by Higgins, the book still has various misprints, and one must be cautious in lifting formulas out of context without checking them.

Bernoulli numbers may be expressed in a variety of ways as combinations of differences of powers of zero (to use the older nomenclature of calculus of finite differences) and hence are combinations of Stirling numbers of the second kind. In the older British literature we could cite numerous papers bearing on the differences of zero, and we would have to mention the work of Boole, Blissard, etc., particularly in the old issues of the Cambridge and Dublin Journal and the Quarterly Journal of Pure and Applied Mathematics, as well as Messenger of Mathematics. However, the list of titles of papers is very lengthy. Vandiver [30] drew up a bibliography which encompasses well over 400 titles of papers dealing with the Bernoulli numbers, and the present writer [18] maintains a card file that shows over 600 entries now. Ultima-

tely what is needed is a kind of modern L. E. Dickson account of the History of the Special Numbers of Bernoulli, Euler, Stirling, Worpitzky, Nielsen, etc.

The problem of information retrieval becomes ever more difficult. It is only because of a maximum of effort to peruse every page in the known journals that the present writer makes any pretense at knowing what has or has not been done with certain of these sequences of numbers in analysis, combinatorics, or number theory. One of the greatest aids in such a retrieval of information has been the review journals and in particular the old *Jahrbuch über die Fortschritte der Mathematik* (1868–1944). This journal is also hard to locate in ordinary libraries and it is fondly hoped that it will be reprinted for the convenience of contemporary young mathematicians who do not have access to such valuable sources of information.

Kronecker [21] discovered an interesting formula for the Bernoulli numbers, which may be stated in the form [24, page 102]

$$B_{2n} = \sum_{j=2}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} \frac{1}{j} \sum_{k=1}^{j-1} k^{2n}. \tag{4}$$

Saalschütz shows the details of derivation from the Lagrange interpolation formula. Recently, Bergmann [1] has essentially rediscovered this formula, though is evidently not aware of this fact. Bergmann stated and proved his formula in the form (proof also by Lagrange formula!)

$$B_{n-1} = -\sum_{j=1}^{n} (-1)^{j} \binom{n}{j} \frac{1}{j} \sum_{k=1}^{j} k^{n-1}. \tag{5}$$

Since the Bernoulli numbers, in our notation, are defined by the expansion

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}, \tag{6}$$

it is known that $B_{2n+1} = 0$ for $n \geqq 1$. Thus we can write Bergmann's formula equivalently as follows:

$$\begin{aligned}
B_{2n} &= \sum_{j=1}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} \frac{1}{j} \sum_{k=1}^{j} k^{2n} \\
&= \sum_{j=2}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} \frac{1}{j} \sum_{k=1}^{j} k^{2n} + (2n+1) \\
&= \sum_{j=2}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} \frac{1}{j} \sum_{k=1}^{j-1} k^{2n} + (2n+1) \\
&\qquad + \sum_{j=2}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} \frac{1}{j} j^{2n} \\
&= \sum_{j=2}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} \frac{1}{j} \sum_{k=1}^{j-1} k^{2n} + \sum_{j=0}^{2n+1} (-1)^{j-1} \binom{2n+1}{j} j^{2n-1},
\end{aligned}$$

and the second sum here is identically zero, being (apart from sign) just a $(2n+1)$th difference of a polynomial of degree less than $2n+1$. Hence Bergmann's formula implies Kronecker's (and the steps are reversible also). This discovery of something equivalent to an old result of Kronecker, and using almost the same proof, points up again the difficulty of finding anything absolutely new. It is also easy to transform (5) directly into (1).

In the author's thesis [15] an attempt was made to unify various formulas pertaining to the Stirling numbers. There extensive use was made of contour integration and generalized chain rule differentiation formulas in order to evaluate the Stirling numbers of the first kind. In passing, the numbers of Worpitzky-Euler-Nielsen were studied. At the time the thesis was written the author had not solved a related problem — to express the Stirling numbers of the second kind in terms of those of first kind (the reverse expansion was quite well known). Later this was solved and published in [16], together with some variations. We shall illustrate the uses of the generalized chain rule by deriving a formula for the Bernoulli numbers. The technique should suggest the general approach that is possible.

First of all, if x is a function of z and all indicated derivatives exist, then the chain rule may be written in the form [15], [25]

$$D_z^n f(x) = \sum_{k=0}^{n} D_x^k f(x) \frac{(-1)^k}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} x^{k-j} D_z^n x^j, \tag{7}$$

where $D_z = d/dz$.

The formula is very useful for determining Taylor coefficients in complicated expansions, and has a number of interesting implications. It may be remarked that this formula has a long and interesting bibliography in and of itself. Among its corollaries are the following formulas:

$$D_z^n x^{-a} = a\binom{a+n}{n} \sum_{j=0}^{n} (-1)^j \binom{n}{j} \frac{1}{a+j} x^{-a-j} D_z^n x^j, \quad a \text{ real}; \tag{8}$$

$$x^a D_z^n x^{-a} = \sum_{j=0}^{n} \binom{-a}{j} \binom{n+a}{n-j} x^{-j} D_z^n x^j; \tag{9}$$

$$D_z^n \left(\frac{1}{x}\right) = \sum_{j=0}^{n} (-1)^j \binom{n+1}{j+1} x^{-j-1} D_z^n x^j. \tag{10}$$

It is this last form which we shall illustrate.

By means of (6) we have, using (10),

$$B_n = D_t^n \left(\frac{t}{e^t-1}\right) \Big|_{t=0} = \sum_{j=0}^{n} (-1)^j \binom{n+1}{j+1} D_t^n \left(\frac{e^t-1}{t}\right)^j \Big|_{t=0},$$

and since it is easily seen (and well known) that

$$(e^t - 1)^j = \sum_{r=j}^{\infty} \frac{t^r}{r!} \sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} k^r,$$

we find readily

$$B_n = \sum_{j=0}^{n} (-1)^j \binom{n+1}{j+1} \frac{n!}{(n+j)!} \sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} k^{n+j}, \tag{11}$$

which formula should be compared to those already quoted. It is not really a new formula, but shows how quickly one may evaluate B_n.

Incidentally, it was shown in [15] that formula (9) can be looked upon as an immediate consequence of the Lagrange interpolation formula. This is due to the fact that

$$\binom{a}{j}\binom{n-a}{n-j} = \prod_{\substack{k=0\\k\neq j}}^{n} \frac{a-k}{j-k}, \quad \text{for } 0 \leqq j \leqq n, \quad a \text{ real.}$$

The numbers of Bernoulli cannot properly be studied apart from the Eulerian numbers. Eulerian numbers ought not to be confounded with the 'Euler' numbers, there being two species of number named after Euler. The Eulerian numbers may be defined by

$$A_{j,n} = \sum_{k=0}^{j} (-1)^k \binom{n+1}{k} (j-k)^n, \tag{12}$$

from which a number of relations follow:

$$A_{j,n} - A_{n-j+1,n} = \begin{cases} 0, & \text{for } n \geqq 1, \\ (-1)^j, & \text{for } n = 0. \end{cases}$$

$$\sum_{j=0}^{n} A_{j,n} = n!, \tag{13}$$

$$x^n = \sum_{j=0}^{n} \binom{x+j-1}{n} A_{j,n} \text{ for all real } x,$$

and so forth. A full discussion of these numbers and their extensions may be found in the papers of Carlitz (and other work of his) as well as in [15]. Most of the numbers studied by Worpitzky, Nielsen, and Euler may be subsumed, as in [15], in the expression

$$B^n_{r,q} = \sum_{k=0}^{r} (-1)^k \binom{q}{k} (r-k)^n, \tag{14}$$

which includes the differences of zero (i.e., Stirling numbers of second kind) because $B^n_{r,r} = \Delta^r 0^n$. These numbers have numerous properties, among which we mention just:

$$B^n_{k,m+1} = (-1)^{m+n} B^n_{m-k+1,m+1} \qquad m \geqq n \geqq 1,$$

$$\sum_{k=0}^{m+1} \binom{k-1}{m-a} B^n_{k,m+1} = (-1)^{m+n} B^n_{a,a} \qquad m \geqq n \geqq 0,$$

$$\sum_{k=0}^{m+1} \binom{x+k-1}{m} B^n_{k,m+1} = (-1)^{m+n} x^n \qquad m \geqq n \geqq 0.$$

We conclude a listing of Bernoulli number formulas by quoting the following from [**15**] and using (14) as a unifying form:

$$B_n = \sum_{k=0}^{n} \frac{(-1)^k}{k+1} B^n_{k,k}, \tag{15}$$

$$B_n = \frac{1}{n+1} \sum_{k=0}^{n} (-1)^k \frac{1}{\binom{n}{k}} B^n_{k,n+1}, \tag{16}$$

$$B_n = \sum_{k=0}^{n} (-1)^k \frac{\binom{n+1}{k+1}}{\binom{n+k}{k}} \frac{1}{k!} B^{n+k}_{k,k}, \tag{17}$$

$$B_{n+1} = \frac{n+1}{2(1-2^{n+1})} \sum_{k=0}^{n} \frac{(-1)^k}{2^k} B^n_{k,k}, \tag{18}$$

$$B_{n+1} = \frac{n+1}{2^{n+1}(2^{n+1}-1)} \sum_{k=0}^{n} (-1)^{k+1} B^n_{k,n+1}, \tag{19}$$

and these do not exhaust the list. Formula (15) is just (1) again which was rediscovered by Higgins. Formula (17) is the same as (11).

It is clear from the results we have cited that far from there being a paucity of ways to express the Bernoulli numbers in closed form as finite sums, quite a lot has been done to obtain such relations. Why then is there so little to be found about these formulas in any of the more commonly used reference works? A variety of reasons may cover the story. A general de-emphasis on technical skills needed for series manipulations is one reason. Recurrence relations have proved of more use in computer calculations than theoretically correct series. The generally increasing volume of mathematical literature is another facet of the information retrieval problem. To English speaking mathematicians it is no deep consolation that a large part of the work on Bernoulli numbers was published in German. The many variant symbols used by mathematicians have also confounded the study of these numbers. These things plus the misinformation prevalent in some quarters that simple formulas for B_n do not exist may explain the situation. Many treatments of the Bernoulli numbers end with a recurrence and some statement to the effect that the numbers

can be expressed in terms of the Riemann zeta function, and indeed we have the well-known formula (not a finite series type)

$$B_{2n} = (-1)^{n-1} \frac{(2n)!}{2^{2n-1}\pi^{2n}} \zeta(2n). \tag{20}$$

Higgins' general formula (2.5) for the Bernoulli numbers, and his result for Euler numbers, are certainly of interest, though the case $a = 0$ in (2.5) yields, as we have seen, a well-known result. What is significant about his general formula is that it ties the Bernoulli numbers in with coefficients of the form

$$\frac{a}{a+bn}\binom{a+bn}{n}$$

by way of the series of Rothe (1793) about which there is a vast literature in and of itself.

We end with a conjecture: the writer has seen no formula for B_n which does not require at least two actual summations. All the formulas we have quoted here are of this type.

References

1. Horst Bergmann, Eine explizite Darstellung der Bernoullischen Zahlen, Math. Nachr., 34 (1967), 377–378.

2. H. Burger, Problem 138, Elemente der Mathematik, 7 (1952), 136–137.

3. L. Carlitz, Generalized Bernoulli and Euler numbers, Duke Math. J., 8 (1941), 585–589.

4. L. Carlitz, q-Bernoulli numbers and polynomials, Duke Math. J., 15 (1948), 987–1000.

5. L. Carlitz, Note on a paper of Shanks, this MONTHLY, 59 (1952), 239–241.

6. L. Carlitz, Remark on a formula for the Bernoulli numbers, Proc. Amer. Math. Soc., 4 (1953), 400–401.

7. L. Carlitz, Expansions of q-Bernoulli numbers, Duke Math. J., 25 (1958), 355–364.

8. L. Carlitz, The Staudt-Clausen theorem, Math. Mag., 34 (1961), 131–146.

9. L. Carlitz, Extended Bernoulli and Eulerian numbers, Duke Math. J., 31 (1964), 667–689.

10. E. Catalan, Sur les différences de 1^p, et sur le calcul des nombres de Bernoulli, Ann. Mat. Pura Appl., 2 (1859), 239–243.

11. E. Cesàro, Transformations algébriques par le calcul des différences, Nouv. Ann. Math., (3)5 (1886), 489–492.

12. G. Frobenius, Über die Bernoulli'schen Zahlen und die Euler'schen Polynome, Sitzungsber. Preuss. Akad. Wiss., (1919), 809–847.

13. H. L. Garabedian, A new formula for the Bernoulli numbers, Bull. Amer. Math. Soc., 46 (1940), 531–533.

14. F. Gomes-Teixeira, Note sur les nombres de Bernoulli, Amer. J. Math., 7 (1885), 288–292.

15. H. W. Gould, The Stirling numbers and generalized difference expansions, Master's Thesis, Univ. of Virginia, 1956.

16. H. W. Gould, Stirling number representation problems, Proc. Amer. Math. Soc., 11 (1960) 447–451.

17. H. W. Gould, Review of O. E. Stanaitis, "An Introduction to Sequences, Series, and Improper Integrals," this MONTHLY 76 (1969), 210–211.

18. H. W. Gould, Bibliography of articles on the special number sequences of Bernoulli, Stirling, Euler, Worpitzky, etc., unpublished card file. About 600 items.

19. James Higgins, Double series for the Bernoulli and Euler numbers, J. London Math. Soc., (2)2 (1970), 722–726.

20. Charles Jordan, Calculus of Finite Differences, Budapest, 1939; Reprinted by Chelsea Publ. Co., New York, 1950, still in print.

21. L. Kronecker, Bemerkung zur Abhandlung des Herrn Worpitzky, J. Reine Angew. Math., 94 (1883), 268–270.

22. Ove J. Munch, Om Potensproduktsummer, Nordisk Mat. Tidsskr., 7 (1959), 5–19.

23. H. Nägelsbach, Zur independente Darstellung der Bernoulli'schen Zahlen, Zeitschr. für Math. und Physik, 19 (1873), 219–234.

24. L. Saalschütz, Vorlesungen über die Bernoulli'schen Zahlen, ihren Zusammenhang mit den Secanten-Coefficienten und ihre wichtigeren Anwendungen, Berlin, 1893. Available since 1964 in Xerographed form from University Microfilms, Ann Arbor, Michigan. Order No. OP-17136.

25. I. J. Schwatt, Introduction to the Operations with Series, Univ. of Pennsylvania Press, 1924; Reprinted by Chelsea Publ. Co., New York, 1962.

26. I. J. Schwatt, Finite expressions for the Bernoulli numbers obtained by the actual expansion of trigonometric functions by Maclaurin's theorem, J. Math. Pures Appl., (9)11 (1932), 143–151.

27. E. B. Shanks, A finite formula for the Bernoulli numbers, this MONTHLY, 59 (1952), 496, Abstract No. 30.

28. K. G. C. von Staudt, De numeris Bernoullianis commentatio, Erlangen, 1845.

29. L. Toscano, Nota bibliographica sui numeri di Stirling di prima specie, Gior. Mat. Battaglini, (6)2(92) (1964), 120–122. Total of 36 items.

30. H. S. Vandiver, Bibliography of articles on Bernoulli and Euler numbers for the years 1869–1940, Mimeographed manuscript of 19 pp. with 4 pp. Addendum covering years 1713–1915. Total of over 400 items.

31. J. Worpitzky, Studien über die Bernoullischen und Eulerschen Zahlen, J. Reine Angew. Math., 94(1883), 203–232.

32. Niels Nielsen, Traité élémentaire de nombres de Bernoulli, Paris, 1923. Chapter 12 is about explicit formulas.

33. N. E. Nörlund, Vorlesungen über Differenzenrechnung, Berlin, 1924; Reprinted by Chelsea Publ. Co., New York, 1954. Explicit formula is used on pp. 32–34 to prove the Staudt-Clausen theorem.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306; notes are usually limited to three printed pages.

A NOTE ON THE MEAN VALUE THEOREM

A. A. GOLDSTEIN, University of Washington

Various generalizations of the ordinary mean value theorem to vector-valued functions are known, [1], [2], $\cdots$, [5]. One of the nicest is a simple result due to McLeod [2]. In applied analysis the upper estimate of $\|f(x)-f(y)\|/\|x-y\|$ by $\sup\{\|f'(x+t(x-y))\|: t\in(0,1)\}$ (Graves [1]), has proven very useful, [4]. A

criterion for a lower positive estimate would also be useful. We show that this can be obtained easily using [**2**].

Let f be a continuous function from a closed line segment $[a,b]$ in a real normed-linear space E to a real normed linear space F. Let M be a countable subset of the open segment (a,b) and assume that the right hand variation

$$f'_+(x) = \lim_{t\to 0+} \left(\frac{f(x+t(b-a))-f(x)}{t\|b-a\|}\right)$$

exists for all $x\in(a,b)\sim M$.

THEOREM. *If for some* $\hat{x}\in(a,b)\sim M$ *and* $q\in(0,1)$,

$$\|f'_+(x)-f'_+(\hat{x})\| \leqq q\,\|f'_+(\hat{x})\|$$

for all $x\in(a,b)\sim M$, *then*

$$(1-q)\,\|f'_+(\hat{x})\|\;\|b-a\| \leqq \|f(b)-f(a)\|.$$

Proof. We denote the convex hull of S by $H(S)$. By Theorem 1 p. 200 of [**2**], $(f(b)-f(a))/\|b-a\|$ belongs to $\overline{H(S)}$, where $S=\{f'_+(x)\colon x\in(a,b)\sim M\}$. Let $g\in F^*$ be chosen such that $\|g\|=1$ and $g(f'_+(x))=\|f'_+(x)\|$. Take any ξ in $[a,b]\sim M$. Then

$$g(f'_+(\xi))=g(f'_+(\hat{x}))+g(f'_+(\xi)-f'_+(\hat{x}))\geqq\|f'_+(\hat{x})\|-\|f'_+(\xi)-f'_+(\hat{x})\|\geqq(1-q)\|f'_+(\hat{x})\|.$$

Thus $g(u)\geqq(1-q)\,\|f'_+(\hat{x})\|$ for all $u\in\overline{H(S)}$, and $g(f(b)-f(a))/\|b-a\| \leqq \|f(b)-f(a)\|/\|b-a\|$.

REMARK. A sufficient condition for the hypothesis of the theorem to hold is that f'_+ satisfy a Lipschitz condition with constant L on $[a,b]\sim M$, and for some $\hat{x}\in[a,b]\sim M$ and $q\in(0,1)$, $\|b-a\|=q\|f'_+(\hat{x})\|/L$. Thus if $L<\infty$, and $\|f'_+(\hat{x})\|>0$, the hypotheses of the theorem can always be satisfied for sufficiently small values of $\|b-a\|$.

Applications.

(1) Let f be a continuous map from $[a,b]\sim M$ to F and suppose that for some $\hat{x}\in[a,b]\sim M$, $\|f(x)-f(\hat{x})\|/\|f(\hat{x})\|\leqq q$ with some $q\in(0,1)$. Then

$$\left\|\int_a^b f(x)\,dx\right\| \geqq \|b-a\|\,(1-q)\,\|f(\hat{x})\|.$$

(2) (Isolation of roots). Suppose f is Gateaux differentiable at x in E, $f(x)=0$, $\|f'(x,h)\|\geqq\mu\|h\|$, and

$$\|f'(x,h)-f'(y,h)\|\leqq L\|h\|\;\|x-y\|$$

for all $h\in E$, some positive μ and L, and $y\in S=\{z\colon\|z-x\|<\mu/L\}$. Then the root x is unique in S.

Supported by the Air Force Office of Scientific Research and the Math. Res. Center, University of Wisconsin.

References

1. L. M. Graves, Taylor's theorem in general analysis, Trans. Amer. Math. Soc., 29 (1927) 163–177.
2. R. M. McLeod, Mean value theorems for vector valued functions, Proc. Edinburgh Math. Soc., 14, Ser II (1964–65) 197–209.
3. J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, 1960.
4. L. V. Kantorovich and C. P. Akilov, Functional Analysis in Normal Spaces, Macmillan, New York, 1964, p. 660.
5. Henri Cartan, Calcul Différentiel, Hermann, Paris, 1967, p. 49.

SOME SHORT PROOFS ON SUBSERIES CONVERGENCE

G. J. O. JAMESON, University of Warwick, England

Let (X,τ) be a topological linear space. We shall say that a series Σx_n is **S-convergent** if $\Sigma_{i=1}^{\infty} x_{n_i}$ is convergent for every increasing sequence (n_i) of integers, and that Σx_n is **BM-convergent** if $\Sigma a(n)x_n$ is convergent for every bounded sequence a of real numbers. (We use the notation $a(n)$ for the nth term of a numerical sequence a.) S-Cauchy and BM-Cauchy series are defined analogously. The set of all subsets of the positive integers will be denoted by Σ, and the set of all finite ones by Φ. When a fixed series Σx_n is under consideration, we shall write $s(\phi)$ for $\Sigma_{n\in\phi} x_n$ (where $\phi \in \Phi$), and when the series is S-convergent, we shall use the analogous notation $s(\sigma)$ for all $\sigma \in \Sigma$. If (X, Y) is a dual pair of linear spaces, we shall use the notation $\langle\ ,\ \rangle$ for the bilinear mapping, and U^0 for the polar of U. The weak topology induced by Y on X will be denoted by $\sigma(Y)$, and the space of continuous linear functionals on (X,τ) by X^*.

We shall give short proofs of the following two theorems on S-convergence:

(1) (Robertson [3]; cf. McArthur [2].) *If Σx_n is S-convergent, then $\{s(\sigma): \sigma \in \Sigma\}$ is compact.*

(2) (The Orlicz-Pettis theorem.) *If (X,τ) is locally convex, and Σx_n is S-convergent with respect to $\sigma(X^*)$, then it is S-convergent with respect to τ.*

At the same time, we shall prove corresponding results for BM-convergence. The common feature of our proofs is the use of continuous images of sets that are known to be compact by Tychonoff's theorem.

THEOREM 1. *If Σx_n is S-convergent, then $\{s(\sigma): \sigma \in \Sigma\}$ is compact.*

Proof. We show that this set is a continuous image of the Cantor space 2^{ω}. An element a of 2^{ω} is a sequence of zeros and ones. Hence we can define

$S(a) = \sum_{i=1}^{\infty} a(i)x_i$; this is simply $s(\sigma)$ for a suitable σ. We need only show that S is continuous. Take a neighborhood U of 0, and a closed neighborhood V such that $V - V \subseteq U$. There exists $\phi_0 \in \Phi$ such that $s(\phi) \in V$ for all $\phi \in \Phi$ disjoint from ϕ_0. (This follows easily from the fact that the series is S-Cauchy.) Since V is closed, $s(\sigma) \in V$ for all $\sigma \in \Sigma$ disjoint from ϕ_0. Suppose that $a, b \in 2^{\omega}$ and $a(i) = b(i)$ for $i \in \phi_0$. Let ϕ_1 be the set of $i \in \phi_0$ for which $a(i) = 1$. Then $S(a) - s(\phi_1) \in V$ and $S(b) - s(\phi_1) \in V$. Hence $S(a) - S(b) \in U$.

The above is valid in a commutative topological group. The question of when the converse is true is discussed in [**3**].

From now on, we assume that (X, τ) is locally convex. The equivalence of the following conditions is elementary (and part of the folklore of the subject):

(i) $\sum x_n$ is S-Cauchy,

(ii) $\sum x_n$ is BM-Cauchy,

(iii) Given a τ-neighborhood U of 0 and $\varepsilon > 0$, there exists N such that

$$\sum_{N+1}^{\infty} |\langle x_n, f\rangle| \leqq \varepsilon \quad \text{for all } f \text{ in } U^0.$$

Let a be an element of I^{ω}, i.e., a real sequence with $|a(i)| \leqq 1$ for each i. If U is a closed, convex, circled τ-neighborhood of 0, and N is as in (iii) (with $\varepsilon = 1$), then it is clear that $\sum_{i=p}^{q} a(i)x_i \in U$ whenever $q > p > N$. Using I^{ω} (with the product topology) instead of 2^{ω}, we can now give the variant of Theorem 1 appropriate to BM-convergence.

THEOREM 2. *Let (X, τ) be a locally convex space, and suppose that $\sum x_n$ is BM-convergent. Then $\{\sum a(i)x_i : a \in I^{\omega}\}$ is compact.*

Proof. For $a \in I^{\omega}$, define $S(a) = \sum_{i=1}^{\infty} a(i)x_i$. It is sufficient to show that S is continuous. Take a neighborhood U of 0, and let V be a symmetric neighborhood such that $V + V + V \subseteq U$. By the above, there exists N such that $\sum_{N+1}^{\infty} a(i)x_i \in V$ for all $a \in I^{\omega}$. Choose a fixed element a of I^{ω}. If $|b(i) - a(i)|$ is sufficiently small for $i \leqq N$, then $\sum_{i=1}^{N} [a(i) - b(i)]x_i \in V$. For such b, we have $S(a) - S(b) \in U$.

If A is a norm-compact subset of l_1, and $\varepsilon > 0$ is given, then there exists N such that $\sum_{r>N} |a(r)| < \varepsilon$ for all $a \in A$. For if

$$G_n = \left\{x \in l_1 : \sum_{r>n} |x(r)| < \varepsilon\right\},$$

then the G_n form an open covering of l_1.

Let F be the set of real sequences that take only a finite number of values. We use the fact that the same subsets of l_1 are compact with respect to $\sigma(F)$ and the norm topology. This does not depend on any sophisticated theorems on weak compactness: it can be proved directly by the "sliding hump" method (e.g., [**1**] p. 284).

THEOREM 3. *Let (X, τ) be a locally convex space. If $\sum x_n$ is S-convergent (or*

BM-convergent) with respect to $\sigma(X^*)$, then it is S-convergent (or BM-convergent) with respect to τ.

Proof. Suppose that Σx_n is S-convergent with respect to $\sigma(X^*)$. Then $\Sigma a(i)x_i$ is $\sigma(X^*)$-convergent for $a \in F$: denote its sum by $S(a)$. It is sufficient to show that Σx_n is S-Cauchy with respect to τ, since a τ-Cauchy sequence that has a $\sigma(X^*)$-limit is τ-convergent to that limit. (The result for BM-convergence also follows.)

For $f \in X^*$, let $T(f)$ be the sequence $\{f(x_n)\}$. Then $T(f) \in l_1$, and for $a \in F$, we have

$$(1) \qquad \langle S(a), f\rangle = \sum_{i=1}^{\infty} a(i)\langle x_i, f\rangle = \langle a, T(f)\rangle.$$

Consequently, T is continuous with respect to the topologies $\sigma(X)$ on X^* and $\sigma(F)$ on l_1. If U is a neighborhood of 0 in X, then U^0 is $\sigma(X)$-compact, so $T(U^0)$ is $\sigma(F)$-compact. By the remarks above, it follows that condition (iii) holds, so that Σx_n is S-Cauchy with respect to τ, as stated.

NOTES. (a) Relation (1) shows that $S^* = T$ in the dual pairs under consideration, and that S is continuous with respect to $\sigma(l_1)$ and $\sigma(X^*)$.

(b) Most earlier proofs of Theorem 3 have used the identity of convergent sequences in $\sigma(F)$ and the norm topology of l_1, but not the identity of compact sets.

(c) The space F, with the topology of pointwise convergence, provides a very simple example of a series that is S-convergent but not BM-convergent, namely Σe_n, where e_n is the sequence having 1 in place n and 0 elsewhere.

References

1. G. Köthe, Topologische lineare Räume I, Berlin, 1960.
2. C. W. McArthur, On a theorem of Orlicz and Pettis, Pacific J. Math., 22 (1967) 297–302.
3. A. P. Robertson, On unconditional convergence in topological vector spaces, Proc. Royal Soc. Edinburgh sect. A, 68 (1969) 145–157.

AN EXPONENTIAL CONGRUENCE OF MAHLER

M. B. NATHANSON, University of Rochester

Let a, b, u, v be nonzero integers with $u > v > 1$. Mahler [2] has shown that the congruence

$$(1) \qquad au^n \equiv b \pmod{v^n}$$

has only finitely many solutions $n > 0$. The proof, using a generalization of the Thue-Siegel theorem on diophantine approximation, is noneffective: it yields no method to compute the solutions of (1). In this note I give an elementary and effective solution of (1) in the case $a = b = 1$.

All letters stand for positive integers. As usual, the greatest common divisor of x

and z is denoted (x,z); $x \mid z$ means x divides z; if p is a prime, $p^m \parallel z$ means $p^m \mid z$ but $p^{m+1} \nmid z$. If f is a positive real-valued function such that $\lim_{n\to\infty} f(n)/2^n = 0$, we write $f = o(2^n)$.

The following lemma is well known [**1**].

LEMMA. *Let p be a prime number, and $u > 1$ an integer not divisible by p. Let y_n denote the order of u in the group of units of the ring Z/p^nZ.*

(i) *If $p > 2$ or if $u \equiv 1 \pmod 4$, let $p^N \parallel (u^{y_1} - 1)$. Then $y_n = y_N$ for $n \leqq N$ and $y_n = y_N p^{n-N}$ for $n \geqq N$.*

(ii) *If $p = 2$ and $u \equiv 3 \pmod 4$, let $2^N \parallel (u^2 - 1)$. Then $y_1 = 1$, $y_n = 2$ for $2 \leqq n \leqq N$, and $y_n = 2^{n-N+1}$ for $n \geqq N$.*

Proof. (i) Clearly, if $p^N \parallel (u^{y_1} - 1)$, then $y_n = y_N$ for $n \leqq N$. I show by induction that for $n \geqq N$, $y_n = y_N p^{n-N}$ and $p^n \parallel (u^{y_n} - 1)$. This is true for $n = N$. Assume true for n. Then $u^{y_n} = p^n r + 1$, where $(r,p) = 1$, and

$$u^{py_n} = (p^n r + 1)^p = 1 + p^{n+1} r + p^{n+2} rs = 1 + p^{n+1} r(1 + ps).$$

Therefore, $p^{n+1} \parallel (u^{py_n} - 1)$ and $y_{n+1} \mid py_n$. But $y_n \mid y_{n+1}$ and $y_n \neq y_{n+1}$ (since, by hypothesis, $p^n \parallel (u^{y_n} - 1)$). Hence, $y_{n+1} = py_n = y_N p^{n-N+1}$.

The argument in case (ii) is similar.

THEOREM. *Let u and v be integers with $u > v > 1$. Let $f = o(2^n)$. Then the congruence*

$$u^t \equiv 1 \pmod{v^n} \tag{2}$$

has only finitely many solutions (n,t) with $t \leqq f(n)$.

Proof. If $(u,v) > 1$, then (2) has no solutions. Suppose that $(u,v) = 1$, and that $p \mid v$, p prime. Then $(u,p) = 1$, and $u^t \equiv 1 \pmod{p^n}$. Let y_n denote the order of u in the group of units of the ring Z/p^nZ; then $y_n \mid t$. If $p > 2$ or $u \equiv 1 \pmod 4$, and if $p^N \parallel (u^{y_1} - 1)$, then, by the lemma, for $n \geqq N$,

$$y_n = y_N p^{n-N} \leqq t \leqq f(n).$$

Since $p^N \leqq u^{y_1} - 1 < u^{y_1} < u^p \leqq u^v$,

$$\frac{2^n}{f(n)} \leqq \frac{p^n}{f(n)} \leqq \frac{p^N}{y_N} \leqq p^N < u^v.$$

Similarly, in the case $p = 2$ and $u \equiv 3 \pmod 4$, if (n,t) is a solution of (2) with $t \leqq f(n)$, then $2^n/f(n) < u^v$. Since $f = o(2^n)$, this inequality is satisfied for only finitely many n, so the number of solutions (n,t) with $n \leqq f(t)$ of (2) is finite.

COROLLARY. *Let u and v be integers with $u > v > 1$. If*

$$u^n \equiv 1 \pmod{v^n} \tag{3}$$

then $2^n/n < u^v$.

Proof. This follows immediately from the proof of the theorem, with $f(n) = n$.

I have not found a similarly elementary and effective solution to the general congruence (1), except in the trivial case when $(u, v) > 1$. (If $p \mid (u, v)$, then (1) implies that $p^n \mid b$. Let $p^N \| b$. Then $n \leqq N$.) In particular, for $n > 1$, solutions of

$$5^n \equiv 2 \pmod{3^n} \tag{4}$$

are unknown. I conjecture that there are none.

References

1. W. J. LeVeque, Topics in Number Theory, vol. I, Addison-Wesley, Reading, Mass., 1956, p. 52.
2. K. Mahler, On the fractional parts of the powers of a rational number, Acta Arith., 3 (1938) 89–93.

ON THE INTEGRAL CUBOID

W. G. SPOHN, Johns Hopkins University

A long-standing problem is whether cuboids (rectangular parallelepipeds) exist for which the edges, face diagonals, and inner diagonals are all integers (Dickson [1, p. 502] and Sierpinski [4, p. 62]). It appears not to have been noted that for the well-known family of solutions yielding integral edges and face diagonals, the inner diagonals cannot be integers.

The problem can be expressed as one of finding solutions in positive integers to the following four equations in seven unknowns:

$$x^2 + y^2 = t^2, \qquad x^2 + z^2 = u^2, \qquad y^2 + z^2 = v^2, \tag{1}$$

$$x^2 + y^2 + z^2 = w^2. \tag{2}$$

Part of the difficulty stems from the fact that the general solution of the system (1) is not known. However, a family of solutions going back to the 18th century (Dickson [1 , p. 497]) is given by

$$x = a(4b^2 - c^2), \qquad y = b(4a^2 - c^2), \qquad z = 4abc \tag{3}$$

for positive integers a, b, c satisfying

$$a^2 + b^2 = c^2. \tag{4}$$

Care must be taken, since x and y may be negative, requiring a sign change. Incidentally, Sierpinski [4, p. 61] slips on this point, saying that solutions of (4) in natural numbers yield solutions of (3) in natural numbers, yet $(a, b, c) = (5, 12, 13)$ gives $(x, y, z) = (2035, -828, 3120)$. This leads to a second slip when he assumes that z has the greatest magnitude, yet $(a, b, c) = (11, 60, 61)$ gives $(x, y, z) = (117469, -194220, 161040)$ and $(a, b, c) = (143, 24, 145)$ gives $(x, y, z) = (-2677103, 1458504, 1990560)$.

One sees that solutions of (3) automatically satisfy the second and third equations in (1), where

$$u = a(4b^2 + c^2), \qquad v = b(4a^2 + c^2), \tag{5}$$

while (4) is needed to establish the first equation in (1).

THEOREM 1. *For x, y, z satisfying* (3) *and* (4), *equation* (2) *is impossible.*

Proof. We have

$$x^2 + y^2 + z^2 = c^2(a^4 + 18a^2b^2 + b^4). \tag{6}$$

The left member however, cannot be a square for positive integers x, y, z since the expression in parentheses is not a square for $ab \neq 0$ (Pocklington [**3**, p. 116]) completing the proof.

The simplest solution of (3) and (4) is given by $(a, b, c) = (3, 4, 5)$, $(x, y, z) = (117, 44, 240)$. There are solutions of (1) of the form (3) not satisfying (4), for example, $(x, y, z) = (-855, 2640, 832)$ for $(a, b, c) = (1, 16, 13)$ and solutions of (1) not of the form (3), for example, $(x, y, z) = (240, 252, 275)$. However, the following theorem demonstrates that (3) in some sense represents all solutions of (1).

THEOREM 2. *Formula* (3) *with* $(a, b, c) = 1$ *represents some integral multiple of every primitive solution of* (1).

Proof. A primitive solution of (1) is one in which x, y, z, t, u, v are positive integers with no common factor. Any solution can be reduced to a primitive solution. In a primitive solution $(x, y, z) = 1$.

For the given x, y, z one solves (3) to get positive real solutions

$$a = \frac{B}{2(AB)^{1/3}}, \qquad b = \frac{A}{2(AB)^{1/3}}, \qquad c = \frac{z}{(AB)^{1/3}}, \tag{7}$$

where .

$$A = x + u, \qquad B = y + v. \tag{8}$$

If a, b, and c are integers, the solution itself is represented. If a, b, and c are not all integers, multiply the given solution by the integer $8AB$ to get the primed solution

$$x' = 8ABx, \qquad y' = 8ABy, \qquad z' = 8ABz, \tag{9}$$

where a', b', c' are integers given by

$$a' = y + v, \qquad b' = x + u, \qquad c' = 2z. \tag{10}$$

One readily sees from (3) that the cube of any common factor of a', b', c' is a factor of x', y', z'. Hence a', b', c' can be reduced to a'', b'', c'', where $(a'', b'', c'') = 1$ and the corresponding x'', y'', z'', t'', u'', v'' is a multiple of the original primitive solution of (1), to end the proof.

Normally, there are 6 such representations from permutations of x, y, z,

though only 3 significant ones, because an interchange of x and y produces an interchange of a and b. This shows that there is no loss in assuming $2a > 2b > c > 0$. If in addition x and y are allowed to be negative, there are 24 representations. Thus a way to check if a particular solution of (1) is represented by (3) would be to examine the 24 sets of solutions for a, b, c given in (7) to see if any set is integral.

Lal and Blundon [2] used the formula

$$x = 2mnpq, \qquad y = mn(p^2 - q^2), \qquad z = pq(m^2 - n^2) \tag{11}$$

to generate solutions of (1), requiring that y^2+z^2 be a square. This also represents multiples, $(x, y, z) = (44, 117, 240)$ not being representable. This fact necessitates that solutions be reduced, and furthermore, makes it difficult to give the range of their table. Though formula (3) also has this defect, it may be more effective, involving one less parameter. A still more preferred form may be

$$x = a(b^2 - c^2), \qquad y = b(a^2 - c^2), \qquad z = 2abc, \tag{12}$$

with $a > b > c > 0$, $(a, b, c) = 1$, and x^2+y^2 a square.

References

1. L. E. Dickson, History of the Theory of Numbers, Vol. 2, Diophantine Analysis, Carnegie Institute of Washington, 1919, Reprint by Chelsea, 1952.
2. M. Lal and W. J. Blundon, Solutions of the Diophantine equations $x^2+y^2=l^2$, $y^2+z^2=m^2$, $z^2+x^2=n^2$, Math. Comp., 20 (1966) 144–147.
3. H. C. Pocklington, Some Diophantine impossibilities, Proc. Cambridge Phil. Soc., 17 (1914) 110–118.
4. W. Sierpinski, Elementary Theory of Numbers, Panstwowe Wydawnictwo Naukowe. Warsaw, 1964.

REFLECTIONS HAVE REVERSED VECTORS

A. M. ADELBERG, Grinnell College

1. Introduction. In this note we prove the following elementary theorem, which gives some geometric insight into the notion of a reflection of a metric vector space:

THEOREM A. *Every reflection has a reversed vector.*

We also show that the preceding theorem is almost immediately equivalent to the following one:

THEOREM B. *Every rotation of a space of odd dimension and every reflection of a space of even dimension has a fixed vector.*

In spite of the elementary nature of these results, we have not been able to locate Theorem A in the literature except in [1] where both theorems are given, but under restrictive hypotheses, namely for real, anisotropic vector spaces. The proof given there rests on properties of the reals, and does not generalize. Theorem B can be found in the literature (see [2], page 131 or [3], Proposition

187.1), but the proofs make essential use of the relatively deep Cartan-Dieudonné Theorem, so it is of some interest to have a direct proof. It will be seen that the judicious use of determinants may substantially simplify many of the arguments on metric vector spaces that occur in the existing literature.

Finally we shall use Theorem B to provide a proof for the special case of the Cartan-Dieudonné Theorem for anisotropic spaces that is very likely briefer than anything currently available.

2. Preliminaries. We recall the main definitions and the basic results that are needed as follows:

Let k be a field of characteristic $\neq 2$, which will be the base field for all vector spaces. If V is a finite-dimensional vector space, V is called a **metric vector space** (mvs) if there is given a symmetric bilinear form $\beta: V\times V\to k$, called the inner product; if X and Y are elements of V, $\beta(X, Y)$ is usually denoted by XY, and in particular $\beta(X, X)$ by X^2.

We say that X and Y are **orthogonal** if $XY=0$, and call X a **null-vector** if $X^2=0$. If $S\subset V$, the **orthogonal complement of** S is $S^*=\{X\in V \mid XY=0$ for all $Y\in S\}$. Clearly S^* is a vector subspace of V. In particular, $V^*=\text{Rad } V$ is a subspace, called the **radical** of V. We say that V is **nonsingular** if Rad $V=\{0\}$. It is possible to show that if V is non-singular and S is a subspace, then dim S $+$dim $S^*=$dim V, hence $S^{**}=S$. V is called **anisotropic** if 0 is the only null-vector. An anisotropic mvs is clearly non-singular, but the converse is not true. An important example of a non-singular mvs which is not anisotropic is the **hyperbolic plane** $V=k^2$, with inner product $(x, y)(x', y')=xx'-yy'$. The null-vectors for this mvs are the vectors (x, y), where $y=\pm x$, so there are 2 null-lines.

If V is a mvs and $\sigma: V\to V$ is a linear isomorphism, σ is called an **isometry** if $(\sigma X)(\sigma Y)=XY$ for all X and Y in V. If $E_1, \cdots, E_n$ is a basis of V, the $n\times n$ matrix $M=(E_iE_j)$ is called the **matrix of the product** with respect to the basis. It is easy to show that V is non-singular if and only if M is, and that σ is an isometry if and only if $(\sigma E_i)(\sigma E_j)=E_iE_j$ for $1\leqq i, j\leqq n$. It follows immediately that if A is the matrix of σ with respect to the given basis, then σ is an isometry if and only if $M=A^tMA$.

Hence if V is non-singular and σ is an isometry, then with the same notations, det $M=$det A^t det M det $A=(\det A)^2$ det M, so det $\sigma=$det $A=\pm 1$. If det $\sigma=1$, then σ is called a **rotation** (proper isometry), while if det $\sigma=-1$, then σ is called a **reflection** (improper isometry). Clearly the set of all rotations is a subgroup of index 2 of the group of all isometries of V.

If $f: V\to V$, a non-zero vector X in V is **fixed** if $f(X)=X$ and **reversed** if $f(X)=-X$. If V is a mvs and U and W are subspaces, we write $V=U\perp W$ when $V=U\oplus W$ and $XY=0$ for all X in U and Y in W, and call V the **orthogonal sum** of U and W. Clearly if σ is an isometry of U and τ is an isometry of W, there is a unique isometry of V extending σ and τ, denoted by $\sigma\perp\tau$. It is not hard to show that if H is a non-singular hyperplane of a non-singular mvs V, then H^* is 1-dimensional, $V=H\perp H^*$, and there is a unique isometry of V called the **symmetry with respect to** H which fixes the vectors in H and reverses the vectors

in H^*, namely $I_H \perp (-I_{H^*})$. Symmetries are clearly reflections (since H^* is a line) and are involutions.

3. Proofs of the theorems. To prove Theorem A, let σ be a reflection of a non-singular mvs V. Then, the notations being as above, to show that σ has a reversed vector we must show that $A+I$ is a singular matrix. But, using basic properties of determinants,

$$\begin{aligned} \det M \det (A + I) &= -\det A^t \det M \det (A + I) = -\det (A^tMA + A^tM) \\ &= -\det (M + A^tM) = -\det (I + A^t)M \\ &= -\det (I + A^t) \det M \\ &= -\det (I + A)^t \det M = -\det (I + A) \det M. \end{aligned}$$

Therefore $2 \det M \det (I+A) = 0$, hence $\det (I+A) = 0$, so $A+I$ is singular.

To demonstrate the equivalence of Theorems A and B, observe that if σ is an isometry of an n-dimensional space, then $-\sigma$ is an isometry, which is a reflection if and only if σ is a rotation and n is odd or σ is a reflection and n is even, i.e., if and only if σ satisfies the hypothesis of Theorem B. Since a vector is fixed for σ if and only if it is reversed for $-\sigma$, the two theorems are clearly equivalent.

4. Remarks. Theorem B can be used to give a quick proof of the special case of the Cartan-Dieudonné Theorem (see [**2**], p. 129) for anisotropic spaces. The Cartan-Dieudonné Theorem, which is important in determining the structure of the group of isometries, says that any isometry of a non-singular n-dimensional mvs is the product of at most n symmetries. If we assume that the mvs V is **anisotropic** then the following inductive proof is legitimate:

For $n=1$, the Cartan-Dieudonné Theorem is obvious since the only isometries of a non-singular line are $\pm I$.

If σ satisfies the hypothesis of Theorem B, then σ has a fixed vector X, which is a fortiori not a null-vector. Letting $V_1 = \langle X \rangle^*$, V_1 is an anisotropic hyperplane of V, and σ induces an isometry σ_1 of V_1. By the induction hypothesis, σ_1 is the product of at most $n-1$ symmetries of V_1. If τ_1 is an isometry of V_1, then $\tau = I_{\langle X \rangle} \perp \tau_1$ is the only isometry of V which fixes X and extends τ_1. If τ_1 is the symmetry with respect to the non-singular hyperplane H_1 of V_1, then τ is the symmetry with respect to $\langle X \rangle \perp H_1$. Hence $\sigma = I_{\langle X \rangle} \perp \sigma_1$ is the product of at most $n-1$ symmetries of V in this case.

If σ does not satisfy the hypothesis of Theorem B, and τ is any symmetry of V, then $\tau\sigma$ does satisfy the hypothesis, hence by the preceding case, $\tau\sigma$ is the product of at most $n-1$ symmetries of V and $\sigma = \tau(\tau\sigma)$ is the product of at most n symmetries (with one chosen arbitrarily).

Attempting to modify the preceding argument for the general case of a non-singular mvs V of dimension n we run into the problem that a fixed vector X may be null, and consequently $\langle X \rangle^*$ may be singular. Thus for the induction step, we would probably need a Cartan-Dieudonné type theorem for singular spaces.

It is possible to adapt the argument to the general case for $n \leqq 3$ as follows:

For $n=1$, non-singular coincides with anisotropic. For $n=2$, it is a nice exercise to show that I is the only isometry of a non-singular plane with a fixed null-vector (the hyperbolic plane is essentially the only example), hence a reflection is a symmetry, and a rotation is the product of 2 symmetries, one being arbitrary. For $n=3$, we show that any isometry with a fixed null-vector is the product of at most 2 symmetries as follows: If σ has a fixed null-vector X, then σ induces an isometry of the singular plane $\langle X\rangle^*$ which leaves X fixed. Hence if $Y\in\langle X\rangle^* - \langle X\rangle$, then Y is non-null (otherwise $\langle X\rangle^*$ would be a null-plane, while in fact $\langle X\rangle = \text{Rad } \langle X\rangle^*$), and $\sigma(Y)=aX\pm Y$ for some $a\in k$. Let τ_1 be the symmetry of V with respect to the non-singular plane $\langle Y\rangle^*$, so that τ_1 fixes X and reverses Y, and τ_2 be the symmetry of V with respect to the non-singular plane $\langle (a/2)X - Y\rangle^*$, so that τ_2 fixes X and reverses $(a/2)X-Y$. We assert that if $\sigma(Y) = aX-Y$, then $\sigma=\tau_2$, while if $\sigma(Y)=aX+Y$, then $\sigma=\tau_1\tau_2$. To see this, let $\sigma'=\tau_2\sigma^{-1}$ in the first case and $\sigma'=\tau_1\tau_2\sigma^{-1}$ in the second. Then σ' is an isometry of V leaving X and Y fixed. σ' induces an isometry of the non-singular plane $\langle Y\rangle^*$ leaving the null-vector X fixed, hence according to the discussion for $n=2$, σ' restricted to $\langle Y\rangle^*$ is the identity. Thus $\sigma'=I$, which completes the proof for $n\leqq 3$.

It is unlikely that this type of elementary argument will suffice to prove the general case of the Cartan-Dieudonné Theorem for $n\geqq 4$ or even for $n=4$, since it is known that there are isometries of a non-singular 4-dimensional mvs with a fixed null-vector, which cannot be represented as the product of fewer than 4 symmetries.

References

1. W. H. Greub, Linear Algebra, 3-rd ed., Springer, New York, 1967, p. 222.
2. E. Artin, Geometric Algebra, Interscience, New York, 1957.
3. E. Snapper and R. Troyer, Affine and Metric Geometry, Notes 1967.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

A PROBLEM CONCERNING SPHERE-PACKINGS AND SPHERE-COVERINGS

L. FEJES TÓTH, Hungarian Academy of Sciences, Budapest

It is well known [1] that the incircles of the regular hexagonal tesselation form a densest circle-packing, and the circumcircles of the tesselation form a thinnest circle-covering of the plane. Thus there is a densest circle-packing and a thinnest circle-covering arising from one another by concentric dilation of the circles.

The situation is quite different in three-space, where both the problem of the densest sphere-packing and the problem of the thinnest sphere-covering are unsolved. It is conjectured that the part of the hexagonal tesselation is taken over by the space-filling of rhombic, or trapezorhombic, dodecahedra in the case of the packing problem, and by the space-filling of truncated octahedra, in the case of the covering problem [1, 2]. Thus it seems highly probable that the sets of centers are completely different in the solutions of the two problems.

One can try to settle the following question without knowing the solution of either of the two basic ones. Prove or disprove the conjecture that in Euclidean 3-space there is no densest packing of congruent spheres such that bigger concentric congruent spheres form a thinnest covering.

The theory of sphere-packing and sphere-covering has a vast literature. References which might be helpful in the solution of this problem are given in [3].

References

1. L. Fejes Tóth, Lagerungen in der Ebene, auf der Kugel und im Raum, Springer-Verlag, Berlin, 1953.
2. ———, Regular figures, Budapest, 1964.
3. C. A. Rogers, Packing and covering, Cambridge, 1964.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306; notes are usually limited to three printed pages.

A NOTE CONCERNING THE SQUARE-FREE INTEGERS

J. E. NYMANN, The University of Texas at El Paso

Throughout this paper S will denote the set of square-free integers and $S(t)$ will denote the number of square-free integers less than or equal to t. It is well known that $S(t) = 6t/\pi^2 + O(\sqrt{t})$. (See, for example, [**1**, p. 83].) The purpose of this note is to obtain this result in, what the author believes to be, a new way and to give an extension of this result for the kth power-free integers. The primary tool for this is the following Moebius inversion formula, which is given for the case $k = 1$ in [**3**, p. 104].

THEOREM 1. *Let f and F be functions defined on $[1, \infty)$. If these functions satisfy the relation*

$$F(t) = \sum_{1 \leqq n^k \leqq t} f(t/n^k), \tag{1}$$

then they also satisfy the "inverse" relation

$$f(t) = \sum_{1 \leqq m^k \leqq t} \mu(m) F(t/m^k), \tag{2}$$

where μ denotes the Moebius function. Conversely, (1) *follows from* (2).

Proof: Assuming (1) holds, we have

$$\begin{aligned} \sum_{1 \leqq m^k \leqq t} \mu(m) F(t/m^k) &= \sum_{1 \leqq m^k \leqq t} \mu(m) \sum_{1 \leqq n^k \leqq t/m^k} f(t/m^k n^k) \\ &= \sum_{\substack{(m,n) \\ 1 \leqq m^k n^k \leqq t}} \mu(m) f(t/(mn)^k). \end{aligned} \tag{3}$$

Here we are summing over all lattice points (m, n), with $m \geqq 1$ and $n \geqq 1$, which lie under the hyperbola $mn = \sqrt[k]{t}$. Now we rearrange the sum collecting terms with $mn = r$, where $1 \leqq r \leqq \sqrt[k]{t}$. Then we obtain from (3)

$$\sum_{1 \leqq r \leqq \sqrt[k]{t}} \sum_{m|r} \mu(m) f(t/r^k) = \sum_{1 \leqq r \leqq \sqrt[k]{t}} f(t/r^k) \sum_{m|r} \mu(m) = f(t),$$

since

$$\sum_{m|r} \mu(m) = \begin{cases} 0 & \text{if } r > 1 \\ 1 & \text{if } r = 1. \end{cases}$$

This derives (2) from (1). The converse is proved by a similar argument.

THEOREM 2. $S(t) = 6t/\pi^2 + O(\sqrt{t})$.

Proof: If n is a positive integer, then $n = r^2 q$, where r is an integer and $q \in S$. Hence

$$[t] = \sum_{1 \leqq r^2 \leqq t} \sum_{1 \leqq q \leqq t/r^2} 1 = \sum_{1 \leqq r^2 \leqq t} S(t/r^2). \tag{4}$$

Applying Theorem 1 to (4), we have

$$\begin{aligned} S(t) &= \sum_{1 \leqq m^2 \leqq t} \mu(m) [t/m^2] = \sum_{1 \leqq m \leqq \sqrt{t}} \mu(m)(t/m^2 + O(1)) \\ &= t \sum_{1 \leqq m \leqq \sqrt{t}} \frac{\mu(m)}{m^2} + O\left(\sum_{1 \leqq m \leqq \sqrt{t}} 1 \right). \end{aligned} \tag{5}$$

The second term of (5) is clearly $O(\sqrt{t})$. For the first term we have

$$\sum_{1 \leqq m \leqq \sqrt{t}} \frac{\mu(m)}{m^2} = \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} - \sum_{m=[\sqrt{t}]+1}^{\infty} \frac{\mu(m)}{m^2}.$$

Now it is well known [2, p. 250] that $\sum_{m=1}^{\infty} (\mu(m))/m^2 = 1/\zeta(2) = 6/\pi^2$, where ζ denotes the Riemann zeta-function. Also

$$\left|\sum_{m=[\sqrt{t}]+1}^{\infty}\frac{\mu(m)}{m^2}\right| < \sum_{m=[\sqrt{t}]}^{\infty}\frac{1}{m^2} < \int_{\sqrt{t}}^{\infty}\frac{dx}{x^2} = \frac{1}{\sqrt{t}}.$$

Hence the first term of (5) is $6t/\pi^2 + O(\sqrt{t})$. The result now follows.

It follows that the natural density of the set of square-free integers is $6/\pi^2$; the **natural density** of S is defined to be $\lim_{n\to\infty} S(n)n^{-1}$ when that limit exists.

By making obvious modifications in the proof of Theorem 2, one can obtain the following generalization.

THEOREM 3. *The number of kth power-free integers less than or equal to t is* $t/\zeta(k) + O(\sqrt[k]{t^{k-1}})$, *where* ζ *denotes the Riemann zeta-function.*

Thus, the natural density of the set of kth power-free integers is $1/\zeta(k)$.

References

1. A. Gioia, The Theory of Numbers, Markham, Chicago, 1970.
2. I. Niven and H. Zuckerman, An Introduction to the Theory of Numbers, Wiley, New York, Second Edition, 1966.
3. H. Rademacher, Lectures on Elementary Number Theory, Blaisdell, New York, 1964.

THE WEIERSTRASS APPROXIMATION THEOREM

EUGENE SCHENKMAN, Purdue University

The object of this note is to construct a certain polynomial and by means of it to give a proof of the Weierstrass approximation theorem.

We begin by noting that

$$\int_0^1 (n+1)(1-y^2)^n 2y\,dy = (1-y^2)^{n+1}\Big|_1^0 = 1.$$

Since $y \leqq 1$ on $[0,1]$, it follows that $\int_0^1 (n+1)(1-y^2)^n 2dy > 1$, and there is a $k_n < n+1$ so that $\int_z^1 k_n(1-y^2)^n 2dy$ is a polynomial $P_n(z)$ such that $P_n(1) = 0$, $P_n(0) = 1$, and (by the symmetry of $(1-y^2)^n$ in the y-axis), $P_n(-1) = 2$. Furthermore $P_n(z)$ is decreasing on $(-1,1)$ since its derivative is positive there.

Now for fixed b with $0 < b < 1$, we have $\lim_{n\to\infty}(n+1)b^n = 0$. It follows that for given $\varepsilon > 0$ and given δ with $0 < \delta < 1$, there is an n so that $(n+1)(1-y^2)^n < \varepsilon/2$ for $\delta \leqq y \leqq 1$, and hence $P_n(\delta) < \varepsilon$. Similarly $P_n(-\delta) > 2 - \varepsilon$. Since $P_n(x)$ is a decreasing function of x on $[-1,1]$, it follows that $P_n(x)$ is between 2 and $2-\varepsilon$ on $[-1,-\delta]$ and is between ε and 0 on $[\delta, 1]$.

Let t be between 0 and 1 and let $Q_n(x) = P_n(x^2 - t)$. Then $Q_n(x)$ is between 2 and $2-\varepsilon$ on $[-h,h]$ with $h^2 = t - \delta$, and $Q_n(x)$ is between 0 and ε on $[-s,-\zeta]$ and $[\zeta, s]$ with $\zeta^2 = t + \delta$ and $\zeta < s < 1$.

Since t, s, and δ are at our disposal, by appropriate change of scale we have the following:

LEMMA. *Let* $0 < a < b < c$ *and let* $m > \varepsilon > 0$. *Then there is a polynomial* $Q(x)$ *which is increasing on* $[-c, 0]$ *and is decreasing on* $[0, c]$, *and so that* $Q(x)$ *is between* $m - \varepsilon$ *and* m *on* $[-a, a]$ *and is between* 0 *and* ε *on* $[-c, -b]$ *and* $[b, c]$.

WEIERSTRASS APPROXIMATION THEOREM. *Let* $f(x)$ *be a real continuous function on a closed interval* $[a, b]$. *Given any* $\varepsilon > 0$, *there is a polynomial* $g(x)$ *so that* $|g(x) - f(x)| < \varepsilon$ *for all* $x \in [a, b]$.

Proof: Suppose without loss of generality that the range of f is $[0, M]$ on $[a, b]$. It will be sufficient to show how to find a polynomial $g_1(x)$ so that $f - g_1$ has range in $[0, .8M]$. For if we can do that, then we can find a polynomial g_2 so that $f - g_1 - g_2$ has range in $[0, .8^2 M]$, and since there is a k so that $.8^k M < \varepsilon$, the polynomial $g = g_1 + g_2 + \cdots + g_k$ is such that $|f(x) - g(x)| < \varepsilon$ for all x in $[a, b]$.

Since f is uniformly continuous on $[a, b]$, there is a k so that if $x, y \in [a, b]$ with $|x - y| < (b - a)/k$, then $|f(x) - f(y)| < .1M$. Let $a = x_0, x_1, \cdots, x_k = b$, be a subdivision of $[a, b]$ so that $x_{i+1} - x_i = (b - a)/k$. Consider the intervals $[x_j, x_{j+1}]$ such that $f(y) \geqq \frac{1}{2}M$ for all $y \in [x_j, x_{j+1}]$, and let P be the set of points of these intervals. Then P is a finite union of closed disjoint (no endpoints in common) intervals $I_1, \cdots, I_r$ such that each I_i is bordered by intervals H_i and J_i of length $(b - a)/2k$, with $f(x)$ in the range $[.4M, .6M]$ on J_i and on H_i.

Now we apply the lemma with $m = \frac{1}{2}M$ and $\varepsilon = M/10r$. For each i there is a polynomial $Q_i(x)$ so that $Q_i(x)$ is between $\frac{1}{2}M$ and $\frac{1}{2}M - \varepsilon$ on I_i and $Q_i(x)$ is between 0 and ε on the complement of $(I_i \cup H_i \cup J_i)$ in $[a, b]$. Then the polynomial $g_1(x) = Q_1(x) + \cdots + Q_r(x) - 2M/10$ has the property that $f - g_1$ has range in $[0, .8M]$, as was to be shown.

WHO DISCOVERED BOYER'S LAW?

H. C. KENNEDY, Providence College

C. B. Boyer [**1**, p. 469] in his recent text, *A History of Mathematics*, has observed: "Clio, the muse of history, often is fickle in the matter of attaching names to theorems!" He was referring particularly to the so-called Maclaurin's Series, noting: "In view of the striking results of Maclaurin in geometry, it is ironic that today his name is recalled almost exclusively in connection with a portion of analysis in which he had been anticipated by some half dozen earlier workers." The observation that theorems are not named after their original discoverers is amply supported in his book, where some thirty such cases are explicitly mentioned in Chapters 18 through 24 (covering, approximately, the period from mid-seventeenth to mid-nineteenth century).

Of course some things have intentionally been named after persons other than their discoverers, such as those named by analogy or relation to another's work. Examples of this are "Peano space" (unknown to Peano), so called because of its connection with Peano's space-filling curve, and the innumerable "Parseval equations", so called because of their similarity to an equation published about 1800 by Marc-Antoine Parseval. There are, however, many instances of mathematical formulas, theorems, etc., which were named after the person thought to have discovered them, only to have an earlier discovery later become known. Examples here are both the Maclaurin and Taylor Series, Picard's Method, and De Morgan's rules in logic. Indeed, Łukasiewicz [3] noted that De Morgan's rules were stated as early as the fourteenth century by William of Ockham [4], and they have recently been found by D. E. Kane [2, p. 180] in the writings of the fifteenth century Paul of Venice.

In recognition of Boyer's statement of this 'law' and his abundant documentation of it, I propose the following:

BOYER'S LAW. *Mathematical formulas and theorems are usually not named after their original discoverers.*

It is perhaps interesting to note that this is probably a rare instance of a law whose statement confirms its own validity!

References

1. C. B. Boyer, A History of Mathematics, Wiley, New York, 1968.
2. D. E. Kane, A Critical Study of the Propositional Logic of Paolo Veneto as seen in his *Logica Magna*, Ph. D. dissertation, River Forest, Illinois, 1970.
3. J. Łukasiewicz, Zur Geschichte der Aussagenlogik, Erkenntnis, 5 (1935–36) 111–131.
4. William of Ockham, Summa Logicae, Pars Secunda, Capitulum 32 [De propositione copulativa].

GALILEO SEQUENCES, A GOOD DANGLING PROBLEM

KENNETH O. MAY, University of Toronto

1. Galileo's idea. In 1615 Galileo observed that the sequence of odd integers had the property

$$\frac{1}{3}=\frac{1+3}{5+7}=\frac{1+3+5}{7+9+11}=\cdots. \tag{1}$$

The observation was closely related to his work on freely falling bodies. Indeed, if distance is proportional to time squared and is one in the first time unit, then the total distances at integral times are the perfect squares, and the incremental distances in successive unit time intervals are the odd integers. If we take a new unit of time equal to some multiple of the original, the ratio of the distances travelled in the first two

time units should be unchanged. But this is just the significance of (1), since it says that the distance in the first n time units is always one third of the distance in the next n time units.

Galileo observed that the sequence of odd integers is the only arithmetic progression with this property, and he considered this an important argument for this law of free fall. Was Galileo right? What can be said about sequences for which the ratio of the sum of the first n terms to the sum of the next n terms is a constant? [1]

2. Dangling problems. This is a typical dangling problem. It can be presented with little symbolism, is easily understood, has intuitive appeal, and is wide open to student initiative in experimenting, formulating questions, conjecturing, and proving. Dangling such a question before a class may lead to general participation in class discussion, group projects, or individual efforts. At the very least it provides the students with a participatory glimse of mathematics in the making. At best it may "turn on" a potential mathematician.

The problem of Galileo sequences was dangled before a class of future teachers at the College of Education at the University of Toronto during 1968–1969. Practically all students participated verbally, and several made significant written contributions [2].

3. Galileo sequences. Let the nth term of a sequence be a_n and the sum of the first n terms S_n. *A Galileo sequence* (*GS*) is a sequence of positive integers satisfying

$$S_{2n} - S_n = pS_n \tag{2}$$

for fixed p ($= 3$ for the odd integers). Equivalent conditions are

$$S_{2n} = qS_n \quad (q = p + 1), \quad \text{and} \tag{3}$$

$$a_{2n-1} + a_{2n} = qa_n. \tag{4}$$

Experimentation suggests many easily proved results relating to sums, differences, multiples, special hypotheses on a_n, etc. In particular:

(5) *If one sequence of positive integers is a multiple of another, then if either is a GS so is the other and they have the same ratios.*

This suggests defining a *primitive GS* as one that is minimal with respect to multiplication. Then it is easy to prove that the only primitive increasing *GS* in arithmetic progression is the odd integers, but that there are many other primitive increasing *GS*.

An early conjecture might be:

(6) *In a GS the second term must be an integral multiple of the first, i.e., p and q are integers.*

To prove this let $q = h/k$ in lowest terms. Then from (4) every a_n is a multiple of

k, and we may form a new GS with the same ratio by dividing all terms by k. Repeating the process with the new GS and its successors m times, we see that k^m divides a_n for arbitrarily large m, which is the case only if $k = 1$.

The most interesting result of the year was the following:

(7) *A necessary and sufficient condition for the existence of a strictly increasing GS is that $p > 2$.*

The following argument is based on the first proof by D. A. Gautreau.

To prove the impossibility for $p = 2$, we show that for any i there is a $j > i$ such that $d_j < d_i$, where $d_n = a_{n+1} - a_n$. Then it follows that eventually the difference of successive terms will be non-positive. In order to prove the inequality, we use the identity

$$d_{2i+1} + 2d_{2i} + d_{2i-1} = 3d_i, \tag{8}$$

which follows from the definition of d_n and (4) with $q = 3$. Now at least one of the three d's in the left member must be less than d_i, for otherwise the left member would be at least $4d_i$.

Since the sequence of odd numbers has $p=3$, we suppose that p is greater than 3, i.e., $p \geqq 4$, $q \geqq 5$.

We claim that a strictly increasing GS is given by $a_1 = 1$,

$$a_{2n-1} = \left[\frac{qa_n - 1}{2}\right], \quad a_{2n} = \left[\frac{qa_n}{2}\right] + 1, \tag{9}$$

where the square bracket indicates the greatest integer function. (The choice is suggested by experiments in which one chooses at each stage the nearest pair of numbers that do not violate the requirements.) Since (9) satisfies (4), and a_{2n-1} is obviously less than a_{2n}, it will be sufficient to prove that $a_{2n} < a_{2n+1}$. This can be done recursively by noting that $a_2 < a_3$ and proving that if $a_n < a_{n+1}$, then $a_{2n} < a_{2n+1}$. From (9) and the fact that $q \geqq 5$,

$$a_{2n+1} - a_{2n} \geqq \frac{qa_{n+1} - 2}{2} - \frac{qa_n + 2}{2} \tag{10}$$

$$\geqq \frac{5}{2}(a_{n+1} - a_n) - 2. \tag{11}$$

But if $a_{n+1} > a_n$, their difference is at least 1 and the right member of (11) is greater than 1/2.

NOTES

1. The problem was suggested by a conversation with Stillman Drake of the Institute for the History and Philosophy of Science and Technology at the University of Toronto. See his *Galileo Studies* (University of Michigan Press, 1970), pp. 218–219, 228.

2. The most substantial contributors were D. A. Gautreau (an auditor from grade 13 of the University of Toronto Schools), S. K. Pasricha, G. C. Reid, F. Riad, and D. Sale. Paul Erdös, while visiting Toronto, concurred in some conjectures under consideration.

PROBLEMS AND SOLUTIONS

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before April 30, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2331. *Proposed by Albert Baake, Sentinel High School, Missoula, Mont.*

Let p be a prime, n a natural number, and let $Z(p^n)$ denote the cyclic group of order p^n. Find all subgroups of a group G which is the direct sum of two copies of $Z(p^n)$.

E 2332. *Proposed by R. S. Luthar, University of Wisconsin*

Find all solutions in positive integers:

$$y^3 + 4y = z^2.$$

E 2333. *Proposed by D. E. Penney, University of Georgia*

If k, m, n are integers, then one solution of the equation

$$\frac{\pi}{4} = k \arctan \frac{m}{n}$$

is $k = m = n = 1$. Find all others.

E 2334. *Proposed by Erwin Just, Bronx Community College*

Let k be an arbitrary positive integer. Prove that there exists a non-integral real number $r > 1$ with the property that k divides $[r^n]$ for every positive integer n. (The square brackets denote the greatest integer function.)

E 2335. *Proposed by J. P. Celenza, Bayside, N.Y.*

Does there exist a continuous function from the reals to the reals which is precisely two-to-one?

E 2336. *Proposed by William Fortney, Dumaguete City, Philippines, and Robert Breusch, Amherst College*

Consider the group of bijective rational functions over the complex numbers (with ∞) under the operation of composition. For any positive integer n, characterize the elements of order n.

SOLUTIONS OF ELEMENTARY PROBLEMS

Telescoping Vandermonde Convolutions

E 2273 [1971, 77]. *Proposed by Øystein Rødseth, University of Bergen, Norway*

Let $\binom{r}{s}$ denote the binomial coefficient with the usual conventions. Prove or disprove the following identity:

$$\sum_{j=1}^{n}\sum_{k=1}^{r} (-1)^{k+1}\binom{m}{k}\binom{r-k+mj-1}{r-k} = \binom{r+mn-1}{r},$$

where m, n and r are positive integers.

I. *Solution by Simeon Reich, Israel Institute of Technology, Haifa.* The following is an instance of a Vandermonde convolution:

$$(1)\qquad \sum_{k=0}^{r} (-1)^{k}\binom{m}{k}\binom{r-k+mj-1}{r-k} = \binom{r+mj-m-1}{r}.$$

(Put $x=-m-1$, $y=mj-1$, and $n=r$ in Formula 3.2 of H. W. Gould, *Combinatorial Identities* (Morgantown, W. Va., 1959).) Rearranging (1) we see that

$$(2)\qquad \sum_{k=1}^{r} (-1)^{k+1}\binom{m}{k}\binom{r-k+mj-1}{r-k} = \binom{r+mj-1}{r} - \binom{r+mj-m-1}{r}.$$

If we sum (2) on both sides from $j=1$ to $j=n$, we see that the right-hand side telescopes and thus

$$\sum_{j=1}^{n}\sum_{k=1}^{r} (-1)^{k+1}\binom{m}{k}\binom{r-k+mj-1}{r-k} = \binom{r+mn-1}{r} - \binom{r-1}{r},$$

which proves the result since $\binom{r-1}{r}=0$.

II. *Solution by M. G. Greening, University of New South Wales, Australia.* The left-hand side of the identity is $(-1)^r$ times the coefficient of x^r in the Maclaurin expansion of

$$F(x) = [1-(1+x)^m]\sum_{j=1}^{n}(1+x)^{-mj}.$$

But an easy computation shows that $F(x)=(1+x)^{-mn}-1$, and the coefficient of x^r in this is

$$\binom{-mn}{r}=(-1)^r\binom{r+mn-1}{r},$$

which proves the result.

Also solved by M. T. Bird, D. M. Bloom, Robert Breusch, H. W. Gould, Robert Heller, Harry Lass, H. W. Soul, J. R. Ventura, M. R. Wise, David Zeitlin, and the proposer.

Composites in a Sequence

E 2274 [1971, 78]. *Proposed by Erwin Just, Bronx Community College*

Let a, t, d and r be arbitrary composite integers with $t\neq 1$. Prove that there exists a set of r consecutive members of the sequence (at^n+d) each of which is composite.

Solution by J. Farrell's Number Theory Class, Butler University. Let $f(n)=at^n+d$; we assume that $(a, d)=(t, d)=1$ since otherwise $f(n)$ is always composite. Note that $3\leqq f(n)<f(n+1)$ for all n. For $i=1, 2, \cdots, r$ let p_i be a prime which divides $f(i)$; then t and any p_i are relatively prime. Now let x_i be the exponent to which t belongs mod p_i and set $x=x_1x_2\cdots x_r$. Then $f(x+i)\equiv f(i)\equiv 0 \pmod{p_i}$ for $i=1, 2, \cdots, r$. Since both $f(i)$ and $f(x+i)$ are multiples of p_i and since $f(x+i)>f(i)$ it follows that $f(x+i)$ is composite for $i=1, 2, \cdots, r$.

Also solved by Irl Bivens, Robert Breusch, S. A. Greenspan, C. V. Heuer & G. A. Heuer, James Long, W. W. Meyer, David Spear, L. J. Warren, W. G. Wild, and the proposer.

A Prime Number Inequality

E 2275 [1971, 78]. *Proposed by R. M. Giuli, San Jose State College*

Let P_k be the kth prime ($P_1=2$). Prove that for $k=1, 2, \cdots$

$$P_k\leqq\frac{k^2+3k+4}{4}\,.$$

Solution by C. V. Heuer, Concordia College. It is known (W. Sierpinski, *Elementary Theory of Numbers*, Warsaw, 1964, p. 150) that $P_k\leqq 36\,k\log k$. If we let $f(k)=(k^2+3k+4)/4$, one easily checks that $36\,k\log k<f(k)$ for $k\geqq 991$. The desired result follows upon checking the first 990 cases: this is easier than it looks, for $P_{176}<\cdots<P_{990}=7829<7877=f(176)<\cdots<f(990)$, so the result is true for all k, $176\leqq k\leqq 990$. Similarly $P_{175}=1039<1040.5=f(63)$, $P_{62}=293<298=f(33)$, etc.

Also solved by D. Borwein & J. M. Borwein, Robert Breusch, J. P. Farrell's Butler University Number Theory Class, S. I. Gendler, Heiko Harborth (Germany), G. L. Isaacs, Lew Kowarski, G. L. Miller, Simeon Reich (Israel), Jonathan Ryshpan, F. G. Schmitt, Jr., R. E. Shafer, L. J. Warren, and A. Zujus.

Isaacs remarks that equality holds only for $P_1=2$ and $P_5=11$. All of the solvers use essentially the same technique—the editors had hoped for a simple proof by induction.

A Gambler's Ruin Problem

E 2276 [1971, 78]. *Proposed by D. M. Bloom, Brooklyn College*

Consider the following game: two players A and B start with n and m counters respectively. At each move, one of these $n+m$ counters is selected at random (with each counter having an equal probability of being selected). Whichever counter is selected changes hands, e.g., if it belonged to A prior to the selection, it belongs to B after. The game continues until one player (the winner) acquires all of the counters. Let $P(n, m)$ be the probability that A wins over B. Prove: for each fixed positive integer m, (a) $\lim_{n\to\infty} P(n, m)=\frac{1}{2}$; (b) $P(n, m)$, considered as a function of n, has its maximum value when $n=m+2$.

Solution to part (a) by F. G. Schmitt, Jr., If X_t denotes the number of counters A has after t moves and if $K=m+n$, then X_t is a random walk on the integers $0, 1, \cdots, K$ with initial state $X_0=n$. The stationary transition probabilities $p_{ij}=Pr\{X_{t+1}=j \mid X_t=i\}$ are given by the following:

$$
\begin{aligned}
&P_{i,i+1} = 1 - i/K && \text{if } i = 1, 2, \cdots, K-1\\
&P_{i,i-1} = i/K && \text{if } i = 1, 2, \cdots, K-1\\
&P_{00} = P_{KK} = 1, && P_{ij} = 0 \text{ otherwise.}
\end{aligned}
$$

(If the barriers 0 and K were reflecting rather than absorbing, then X_t would be the random walk corresponding to the Ehrenfest urn model of diffusion.) Let us write $f_n \equiv p(n, K-n)$ for the absorption probability $f_n=\Pr\{X_t=K \text{ for some } t \mid X_0=n\}$. These quantities satisfy the following equations

$$
\begin{aligned}
&f_n = (n/K)f_{n-1} + (1 - n/K)f_{n+1} \qquad \text{for } n = 1, 2, \cdots, K-1\\
&f_0 = 0, \qquad f_K = 1.
\end{aligned}
$$

Rewriting this as

$$f_{n+1} - f_n = \frac{n}{K-n}(f_n - f_{n-1}),$$

recursion yields

$$f_{n+1} - f_n = \frac{n!}{(K-1)_n}(f_1 - f_0) = \binom{K-1}{n}^{-1} f_1.$$

Hence, for $n=1, 2, \cdots, K$,

$$f_n = \sum_{j=0}^{n-1}(f_{j+1} - f_j) = f_1 \sum_{j=0}^{n-1}\binom{K-1}{j}^{-1}.$$

For $n=K$ this becomes

$$1 = f_K = f_1 \sum_{j=0}^{K-1}\binom{K-1}{j}^{-1},$$

so that

$$f_n = \left\{\sum_{j=0}^{n-1} \binom{K-1}{j}^{-1}\right\} \div \left\{\sum_{j=0}^{K-1} \binom{K-1}{j}^{-1}\right\}.$$

Therefore

$$P(n, m) = \frac{\sum_{j=0}^{n-1} \binom{m+n-1}{j}^{-1}}{\sum_{j=0}^{m+n-1} \binom{m+n-1}{j}^{-1}} = \frac{\sum_{j=0}^{n-1} j!(m+n-1-j)!}{\sum_{j=0}^{m+n-1} j!(m+n-1-j)!}.$$

Now, let $n \to \infty$ so that, in particular, $n > m$. Then

$$P(n, m) = \frac{\sum_0^{m-1} + \sum_m^{n-1}}{\sum_0^{m-1} + \sum_m^{n-1} + \sum_n^{m+n-1}};$$

but $\sum_0^{m-1} = \sum_n^{m+n-1}$, so that

$$\begin{aligned} P(n, m) &= 1 - \frac{\sum_0^{m-1}}{2\sum_0^{m-1} + \sum_m^{n-1}} \\ &= 1 - \frac{\sum_0^{m-1} j!(m+n-1-j)!}{O((n-1)!) + 2\sum_0^{m-1} j!(m+n-1-j)!} \\ &= 1 - \frac{1}{2 + O(1/n)} = \tfrac{1}{2} + O(1/n). \end{aligned}$$

Also solved (part (a)) by Ellen Hertz, Harry Lass, J. M. Reiner, and the proposer.

Editorial Comment. The proposer, using a very complicated argument which we must omit for lack of space, shows that

$$\begin{gathered} P(0, m) < P(1, m) < \cdots < P(m+1, m) < P(m+2, m), \\ P(m+2, m) > P(m+3, m) > \cdots \end{gathered}$$

for all positive m with the single exception $P(3, 1) = P(4, 1)$. Lass obtains the partial result $P(m+1, m) < P(m+2, m)$, $P(m+2, m) > P(m+3, m) > \cdots$, again with the exception noted above.

Reiner remarks that the problem is a special case of the Ehrenfest urn problem (Phys. Zeit. 8 (1907), 311–314) which has been treated, among others, by Mark Kac (this MONTHLY, 54 (1947), 363–391).

C(rook)ed Paths

E 2278 [1971, 196]. *Proposed by Henry Cheng, University of California, San Diego*

What is the number of shortest paths from one corner of a chessboard to the diagonally opposite corner which can be traversed by a rook in seven moves, but no fewer?

Solution by Jordi Dou, Barcelona, Spain. Any suitable path consists of three segments totalling seven squares parallel to one edge, alternating with four segments totalling seven squares perpendicular to that edge. The number of partitions of 7 into three positive integers is $\binom{6}{2} = 15$, and the number of partitions of 7 into four positive integers is $\binom{6}{3} = 20$. Thus the number of paths with three horizontal segments is $15 \cdot 20 = 300$; there are 300 more paths with three vertical segments, for a total of 600 paths.

Also solved by P. H. Anderson, R. M. Anderson, Walter Bluger, R. L. Breusch, Butler University Number Theory Class, Cal Poly Solution Group, R. B. Davis, R. L. Enison, Neal Felsinger, E. T. Frankel, J. K. Gendler, Heiko Harborth (Germany), C. V. Heuer, M. Hirschhorn (Scotland), J. C. Hudson, Carolyn MacDonald, Robert Patenaude, Paul Payne, K. R. Rebman, H. S. Sun, R. K. Tamaki, W. G. Wild, and the proposer.

Harborth and Patenaude consider the following generalization: What is the number of shortest paths from one corner of an $m \times n$ chessboard to the diagonally opposite corner which can be traversed by a rook in k moves, but no fewer? By an argument analogous to Dou's they show that the answer is

$$\binom{m-2}{s} \cdot \binom{n-2}{t} + \binom{m-2}{t} \cdot \binom{n-2}{s}, \text{ where } s = [\tfrac{1}{2}(k-1)] \text{ and } t = [\tfrac{1}{2}(k-2)].$$

Tamaki refers to similar problems in Feller, *An Introduction to Probability Theory and its Applications*, 3rd edition, 1968, Vol. I, p. 38, and notes that in the related problem on p. 36 the answer given as $\binom{16}{8}$ should in fact be $\binom{14}{7}$.

Several incorrect solutions were received. One possible misunderstanding came from the fact that the restriction "seven moves but no fewer" implies that there cannot be two successive moves in the same direction, since this would mean that the two could be accomplished as one.

Consecutive Composite Numbers

E 2279 [1971, 196]. *Proposed by Erwin Just, Bronx Community College*

It has been shown (see C. A. Grimm, *A conjecture on consecutive composite numbers*, this MONTHLY, 76 (1969) 1126–1128) that each member of the sequence of integers, $n! + 2, n! + 3, \cdots, n! + n$, is divisible by a prime which does not divide any other member of the sequence. Prove that for any positive integers n and k, there exists a sequence of n consecutive integers such that each member of this sequence is divisible by k distinct prime factors no one of which divides any other member of the sequence.

I. *Solution by David Spear, City College, New York.* For $n = 1$ the result is trivially true, so let n and k be given with $n \geqq 2$. Choose nk distinct primes q_{ij} ($i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, k$) such that $q_{ij} > n$ for all i, j. For $i = 1, 2, \cdots, n$ set $M_i = q_{i1}q_{i2} \cdots q_{ik}$. Then the M_i are pairwise relatively prime so that the Chinese Remainder Theorem guarantees solutions of the following system of simultaneous congruences: $x + i \equiv 0 \pmod{M_i}$, $i = 1, 2, \cdots, n$. If x is any solution, then $M_i \mid x + i$ so that $x + 1, x + 2, \cdots, x + n$ is a sequence of the required type. Note that since each $q_{ij} > n$, each q_{ij} divides one and only one term of the sequence.

II. *Comment by C. A. Grimm, South Dakota School of Mines and Technology.* This problem is a special case of a Classroom Note of mine (this MONTHLY 68 (1961), p. 781). One has only to take the c_j of my note to be a product of k primes, different for each c_j, and each prime greater than n. One should note that the number of primes for each c_j can be varied.

A more general result can be proved along the same lines. Let R be an infinite Euclidean ring and suppose that $p_1, p_2, \cdots, p_n$ are pairwise relatively prime elements of R. Let $m_1, m_2, \cdots, m_{n-1}$ be arbitrary elements of R. Then there exist infinitely many sequences $(c_1, c_2, \cdots, c_n)$ of elements of R such that $p_i \mid c_i$ for $i = 1, 2, \cdots, n$ and such that $|c_{i+1} - c_i| = |m_i|$ for $i = 1, 2, \cdots, n - 1$.

Also solved by Walter Bluger, Butler University Number Theory Class, Cal Poly Solution Group, S. I. Gendler, Heiko Harborth (Germany), C. V. Heuer, Alfred Kohler, Arthur Marshall, Simeon Reich (Israel), and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before April 30, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

5832. *Proposed by Erwin Just, Bronx Community College*

Let n be an integer greater than one. Must there exist an algebraic real number, r, of degree n such that for each positive integer, m, $[r^m]$ is an odd integer?

5833. *Proposed by J. Bernard and G. Letac, University of Clermont-Ferrand, Aubière, France*

Let f be a continuous function on the positive real numbers such that $f(x) \leqq f(nx)$ for any $x > 0$ and any integer $n > 0$. Prove that $\lim_{x\to\infty} f(x)$ exists ($\leqq +\infty$).

5834. *Proposed by Erwin Just, Bronx Community College*

Let f be an irreducible seventh degree polynomial with rational coefficients, and let S be a proper subset of the zeros of f. Can the sum of the elements of S be rational?

5835. *Proposed by G. Letac, University of Clermont-Ferrand, Aubière, France*

Prove that the constants are the only measurable functions f on the positive real line such that for any positive x and $y, f(x+y)$ belongs to the interval spanned by $f(x)$ and $f(y)$.

5836. *Proposed by Eric Bedford and Michael Taylor, University of Michigan*

Let $f(x)$ be bounded and measurable on $(0,1)$. Is it true that $\lim_{n\to\infty} f(x-1/n) = f(x)$ almost everywhere? Prove, or provide a counterexample.

5837. *Proposed by I. N. Herstein, University of Chicago, and Susan Montgomery, University of Southern California*

A theorem of Marshall Osborn states: *If R is a simple ring of characteristic not 2 with an involution such that every non-zero symmetric element is invertible, then either R is a division ring or is 4-dimensional over its center.* Show that if R is a prime ring with involution. of characteristic 2, and if every non-zero symmetric element of R is invertible, then R must be a division ring.

SOLUTIONS OF ADVANCED PROBLEMS

Irreducible Representations of Degree 2 of Simple Groups

5769 [1970, 1115]. *Proposed by L. W. Shapiro, Howard University*

Show that a finite simple group has no irreducible representation over the complex numbers of degree two.

Solution by D. M. Bloom, Brooklyn College. Suppose the finite simple group G has such a representation F. Since deg $F>1$, G is non-abelian. Since G is simple, F is faithful and hence we may regard G as a subgroup of the non-singular 2×2 matrices over C. The set $S=\{A\in G:\det A=1\}$ is a normal subgroup of G; thus $S=\{I\}$ or G. If $S=\{I\}$, then $A\to\det A$ is a monomorphism of G into the *abelian* group C^*; hence $S=G$. Since $-I$ is the only 2×2 matrix of order 2 over C which has determinant 1, and since G has even order (being simple), it follows that $-I\in G$. But the scalar matrices in G form a non-trivial normal subgroup which is abelian (and hence proper) so that G is not simple.

Also solved by I. K. Abroub, L. J. Alex, P. R. Chernoff, E. R. Gentile & M. I. Krusemeyer (Netherlands), M. G. Greening (Australia), J. E. Humphreys, A. A. Jagers, Peter Landweber, Forrest Richen, R. L. Roth, Sister Janet Schillinger, W. C. Waterhouse, and the proposer.

Unique Fixed Point in a Complete Metric Space

5775 [1971, 84]. *Proposed by Simeon Reich, Israel Institute of Technology, Haifa*

Let X be a complete metric space with metric d, let $T:X\to X$, and let $t:X\to$

Reals be defined by $t(x)=d(x, T(x))$. Suppose (1) t is lower-semicontinuous, (2) there exists a sequence $\{x_n\}\subset X$ such that $t(x_n)\to 0$, and (3) $d(T(x), T(y)) \leqq at(x)+bt(y)+cd(x, y)$ where a, b, c are nonnegative, $c<1$ and $x, y\in X$. Prove that T has a unique fixed point and, further, that no one of the three stated conditions can be omitted.

Solution by D. G. Belanger, University of South Alabama. The third condition implies that any sequence $\{x_n\}$ with $\{t(x_n)\}\to 0$ has a convergent subsequence. Using the triangle inequality and (3) we obtain

$$d(x_n, x_m) \leqq t(x_n) + t(x_m) + d(T(x_n), T(x_m)),$$
$$d(x_n, x_m) \leqq (1 + a)t(x_n) + (1 + b)t(x_m) + cd(x_n, x_m).$$

Eventually

$$t(x_n) \leqq \frac{(1 - c)\epsilon}{2(1 + a)} \quad \text{and} \quad t(x_m) \leqq \frac{(1 - c)\epsilon}{2(1 + b)},$$

where $\epsilon>0$. Thus $\{x_n\}$ is Cauchy and has a subsequence $\{x_j\}\to x\in X$. Since t is lower-semicontinuous, $t(x)=0$ and $T(x)=x$.

Let x and y be fixed points in X and $x\neq y$; then

$$0 \neq d(x, y) \leqq d(T(x), T(y)) \leqq cd(x, y).$$

This contradicts $c<1$, hence there is a unique fixed point.

The following examples on R^1 demonstrate that conditions (1), (2), and (3), respectively, are necessary.

(a) Let a, b, and c be arbitrary; $T(x) = \begin{cases} \frac{1}{2}c\,|\,x\,| & \text{if } x \neq 0, \\ \frac{1}{2}c & \text{if } x = 0. \end{cases}$

(b) Let $a = b = 2$; $T(x) = \begin{cases} 1 & \text{if } x < 0, \\ -1 & \text{if } x \geqq 0. \end{cases}$

(c) Let a, b, and c ($\geqq 1$) be arbitrary; $T(x)=\sqrt{2+x^2}$; (it can be proved in this case that $d(T(x), T(y))/d(x, y)\to 1$ as $(x, y)\to\infty$).

Also solved by K. F. Andersen, G. F. Battle, D. F. Behan, P. R. Chernoff, D. K. Cohoon, R. J. Driscoll, Joe Flowers, Hal Forsey, A. A. Jagers (Netherlands), Emmett Keeler, J. R. Kuttler, H.-E. Lahmann, O. P. Lossers (Netherlands), Beatriz Margolis (Argentina), P. J. Owens (England), K. H. Price & C. W. Proctor, Walter Read, J. L. Solomon, E. Y. State, and the proposer.

In the counterexample (c) above, there is no fixed point. Several solvers use the identity transformation in which every point is fixed. Jagers shows that hypothesis (1) may be omitted if a or b is less than 1; (2) may be omitted if $a+b+c<1$.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 2

CONTENTS

(*Continued on inside cover*)

FEBRUARY 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the Monthly. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to Harley Flanders, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

Published by the Association at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

the three-year period 1968–70, is the twentieth award of the Chauvenet Prize since its institution by the MAA in 1925. For the list of the names of previous winners, see this MONTHLY, 71(1964), p. 589, 72 (1965), pp. 2–3, 74 (1967), p. 3, 75 (1968), pp. 3–4, 77(1970), pp. 117–118, and 78 (1971), pp. 112–113.

Professor Trèves was born on April 23, 1930, in Brussels, Belgium. He received the first and second Baccalaureate degrees in Paris in 1949 and 1950, his *licence en science* and his Ph. D. at the Sorbonne in 1953 and 1958. From 1958 to 1961, he was an assistant professor at the University of California, Berkeley, from 1961 to 1964 an associate professor at Yeshiva University, and from 1964 to 1970 a professor at Purdue University. Since 1970, he has been a professor at Rutgers University.

Professor Trèves was an Alfred P. Sloan Fellow in 1960–62 and 1962–64. From June to November 1961 he was under the auspices of the Organization of American States at the *Instituto de Matematica Pura e Aplicada in Rio de Janeiro, Brazil*; in September 1965, he was a Visiting Professor at the Tata Institute of Fundamental Research in Bombay, India, and from 1965 to 1967, and again from May to June, 1970, he was a Visiting Professor at the Sorbonne in Paris.

Professor Trèves' significant contributions to various branches of analysis, but, in particular, to partial differential equations and functional analysis, are contained in his sixty publications.

In accepting the Award, Professor Trèves stressed that he was very much honored and thankful for having been awarded the 1972 Cauvenet Prize. He added that, because of the apparent increasing technicality of mathematical research, it is becoming ever more difficult to exchange information between mathematicians working in different fields — or even in the same field. He felt this to be a worrisome situation, which makes expository talks and articles more necessary than ever.

CONJECTURES AND COUNTEREXAMPLES IN METRIZATION THEORY

L. A. STEEN, St. Olaf College

Prologue. The search for necessary and sufficient conditions for the metrizability of topological spaces is one of the oldest and most productive problems of point set topology. Alexandroff and Urysohn [4] provided one solution as early as 1923 by imposing special conditions on a sequence of open coverings. Nearly ten years later R. L. Moore chose to begin his classic text on the Foundations of Point Set Theory [41] with an axiom structure which was a slight variation of the Alexandroff and Ury-

Lynn Steen received his M. I. T. Ph. D. in 1965 under Kenneth Hoffman and has been at St. Olaf College except for sabbatical leave in 1970–71 at the Inst. Mittag-Leffler, Sweden. His research is centered on topology, and he is the author with J. A. Seebach, Jr. of *Counterexamples in Topology* (HRW, 1970). *Editor.*

sohn metrizability conditions. After Jones **[28]**, we now call any space which satisfies Axiom 0 and parts 1, 2, 3 of Axiom 1 of **[41]** a Moore space. Each metric space is a Moore space, but not conversely, so the search for a metrization theorem became that of determining precisely which Moore spaces are metrizable. The most famous conjecture was that each normal Moore space is metrizable.

It would probably be no exaggeration to say that for the last 30 years, the normal Moore space conjecture dominated the search for a significant metrization theorem and in the process played a major role in the development of point set topology. The conjecture itself was first stated in 1937 by Jones **[28]** who showed that if $2^{\aleph_0} < 2^{\aleph_1}$, then every separable normal Moore space is metrizable. The next major result came nearly twenty years later when Bing **[10]** and Nagami **[44]** showed that every paracompact Moore space is metrizable. But Jones' result together with more recent ones of Heath **[26]** and Bing **[8]** indicated a close relationship between the normal Moore space conjecture and the continuum hypothesis which was shown by Cohen **[18]** in 1963 to be independent of the axioms of set theory. Quite recently Tall and Silver **[54]** used a Cohen model to show that the normal Moore space conjecture itself could not be proved from the present axioms of set theory.

Thus as metrization research shifts from topology to logic, we survey in this paper the chief topological milestones of the last half century. We shall not present proofs that are available in the literature, but shall concentrate instead on gathering together the most significant definitions, theorems, conjectures and counterexamples. The latter will be grouped together at the end of the paper and referenced throughout the text whenever appropriate. We begin at the beginning.

Basic Definitions. We shall assume throughout this paper that all topological spaces are Hausdorff. Most often we shall be concerned only with regular spaces, though this assumption will not go unwritten. **Regular** spaces are those which admit a separation of a point from a closed set by disjoint open neighborhoods. A space X is **normal** if each pair of disjoint closed sets can be separated by disjoint open neighborhoods, and **completely normal** if the same can be done for separated sets. A space is completely normal if and only if it is **hereditarily normal** **[21]**, that is, if and only if every subspace is normal.

A subset of a topological space which can be written as the countable union of closed sets is called an F_σ**-set**; the complement of an F_σ-set can be written as a countable intersection of open sets, and is called a G_δ**-set** (or an **inner limiting** set). A space in which every closed set is G_δ (or equivalently, every open set is F_σ) will be called a G_δ**-space**; a normal space which is also a G_δ-space is called (by Cech **[15]**) **perfectly normal.** Every metric space is perfectly normal and every perfectly normal space is completely normal [33], so we have the following implications:

$$\text{Metrizable} \Rightarrow \text{perfectly normal} \Rightarrow \text{completely normal} \Rightarrow \text{normal} \Rightarrow \text{regular}.$$

Examples 5, 2, 10, and 6 show that none of these implications is reversible.

If a topological space has a countable dense subset it is called **separable**, if it has a countable basis it is **perfectly separable** (or **second countable**), and if it has a countable local basis at each point it is *first countable*. A space in which every subspace is separable is called **hereditarily separable.** If every open covering of X has a countable subcovering, X is called **Lindelöf** (or, by Russian mathematicians, **finally compact** [**3**]); clearly each perfectly separable space is both Lindelöf and hereditarily separable.

Since in a metric space the (open) balls of radius $1/n$ form a countable local basis at each point, every metric space is first countable. Metric spaces need not be second countable, but in metric spaces the properties of separable, hereditarily separable, second countable and Lindelöf coincide. Urysohn [**60**] proved in 1925 that every normal second countable space is metrizable, and, in response to a question proposed by Urysohn, Tychonoff [**59**] showed a year later that every regular second countable space is metrizable.

Developments. A collection of sets $F = \{U_\alpha\}$ is said to **cover** a space X if each point of X belongs to some U_α; if each U_α is open, the cover F is called an **open covering** of X. A cover $\{V_\beta\}$ of a space X is a **refinement** of a cover $\{U_\alpha\}$ if for each V_β there is a U_α such that $U_\alpha \subset V_\beta$. If $S \subset X$, the **star** of S with respect to a cover $F = \{U_\alpha\}$ is the union of all sets in F which intersect S; the star of S is denoted by $F^*(S)$, and the star of the singleton $\{x\}$ is usually denoted simply by $F^*(x)$.

A **development** for a topological space X is a countable family $\mathscr{F}$ of open coverings F_i such that if C is a closed subset of X and $p \in X - C$, there is a covering $F \in \mathscr{F}$ such that no element of F which contains p intersects C (i.e., such that $F^*(p) \cap C = \varnothing$). A space with a development is called **developable.** If $\mathscr{F} = \{F_i\}$ is a development where $F_i \subset F_{i+1}$ for all i, the family $\mathscr{F}$ is called a **nested development**, and if F_{i+1} is a refinement of F_i, $\mathscr{F}$ is called a **refined development**. Clearly each nested development is a refined development; Vickery [**61**] showed that every developable space has a nested development. Axiom 0 and parts 1, 2, and 3 of Axiom 1 of Moore [**41**] require precisely that a space be regular with a nested development $\{F_i\}$; such spaces are called Moore spaces (after Jones [**28**]), and are characterized by the fact that for each $p \in X$, $\{F_i^*(p)\}$ is a countable local basis. Vickery's theorem can be restated as follows: a topological space is a Moore space if and only if it is regular and developable.

Each metric space is a Moore space since the sequence of open coverings by metric balls of radius $1/n$ is a development; examples 6, 9, 14, and 15 show that Moore spaces need not be metrizable.

Semimetric Spaces. A **semimetric** for a Hausdorff space X is a symmetric function $d : X \times X \to R^+$ such that $d(x,y) = 0$ if and only if $x = y$, and if $x \in X$ and $E \subset X$, $\inf \{d(x,y) \mid y \in E\} = 0$ if and only if $x \in \bar{E}$, the closure of E; a Hausdorff space which admits a semimetric is called a **semimetric space.** If we did not require d to be symmetric, to assert the existence of a function with the remaining properties would

be equivalent to saying that the space X was first countable [**13**]. Thus a semimetric space may be thought of as a symmetric first countable space. In fact, some Russian mathematicians call these spaces **symmetrizable**.

Now every developable space has a natural semimetric: if $\{F_n\}$ is a nested development for X (with $X \in F_1$), we define $d(x,y) = \inf \{1/n \mid x,y \in U \in F_n\}$. Then d is a semimetric, but clearly not a metric since d is not continuous. (A semimetric space is metrizable if and only if it has a continuous semimetric [**13**].) Semimetric spaces share with metric spaces the property that every closed set is a G_δ [**35**], hence such spaces are G_δ-spaces. We use Figure 1 to summarize the implications for regular spaces; counterexamples to the converse implications are listed below each implication arrow.

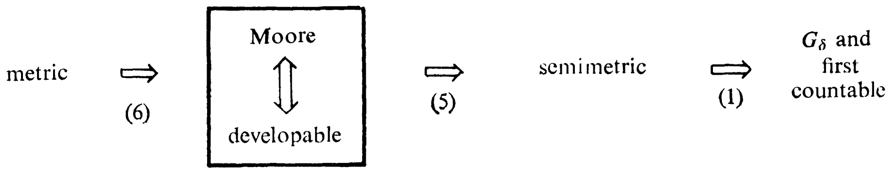


FIG. 1

Every known example of a Moore space which is not metrizable is also not normal; the normal Moore space conjecture asserts that it will always be thus. Jones [**28**] in 1937 mounted the first major attack on this conjecture, and succeed only in proving several weaker theorems: every normal Moore space is completely normal, and every separable normal Moore space is metrizable provided $2^{\aleph_1} > 2^{\aleph_0}$ — a fact implied by (but not equivalent to) the continuum hypothesis. Both of Jones' results have recently been strengthened: McAuley [**36**] observed in 1954 that a simple modification of Jones' proof will show that every normal semimetric space is completely normal, while in 1964 Heath [**25**] showed that a necessary and sufficient condition for the metrizability of a separable Moore space is that every uncountable subset M of the real line contains a subset which is not F_σ (in M). This condition is (perhaps not strictly [**25**]) weaker than that used by Jones, namely $2^{\aleph_0} < 2^{\aleph_1}$.

Jones actually showed that if $2^{\aleph_0} < 2^{\aleph_1}$, then every separable normal space has the property that every uncountable subset has a limit point; Heath [**26**] called spaces with this property **$\aleph_1$-compact** and proved the converse to Jones' theorem: if every separable normal space is $\aleph_1$-compact, then $2^{\aleph_0} < 2^{\aleph_1}$.

Paracompactness. The most significant general approximation to the normal Moore space conjecture is the Bing-Nagami theorem that every paracompact Moore space is metrizable. To develop the concept of paracompactness and all its variations, we must first discuss the naming of various covers.

A cover is **point finite** if each point belongs to only finitely many sets in F, **locally finite** if each point has some neighborhood which intersects only finitely many members of F, and **star finite** if each set in F intersects only a finite number of other sets in F. A cover $V = \{V_\beta\}$ of X is a **star refinement** (or a **point star refinement**, or a Δ **refinement**) of a cover $\{U_\alpha\}$ if for each $x \in X$ there is some U_α such that $V^*(x) \subset U_\alpha$ (where $V^*(x)$ is the star of x with respect to $V = \{V_\beta\}$).

A Hausdorff space is called **fully normal** if every open cover has an open star refinement, **strongly paracompact** (or **star paracompact**) if every open cover has an open star finite refinement, **paracompact** if every open cover has an open locally finite refinement, and **metacompact** (or **pointwise paracompact**, or **weakly paracompact**) if every open cover has an open point finite refinement.

Fully normal spaces were first defined by Tukey **[58]** in 1940, while paracompact spaces were introduced by Dieudonné **[19]** in 1944. Tukey showed that every metrizable space is fully normal, while Dieudonné showed that every paracompact space is normal. The key link between these definitions was provided by Stone **[53]** in 1948 who showed that every metric space is paracompact by proving that every fully normal space is paracompact, and conversely. Although a regular semimetric space need not be paracompact (Example 6), Ceder **[16]** showed that each regular hereditarily separable semimetric space is paracompact. Smirnov **[48]** showed that a paracompact space which fails to be metrizable must fail for local reasons: every locally metrizable paracompact space is metrizable.

Also in 1948 Morita **[43]** introduced the concept (but not the name) of strongly paracompact spaces; he showed that each regular Lindelöf space is strongly paracompact while every strongly paracompact space is *a fortiori* paracompact. Kaplan **[32]** and Alexandroff **[1]** showed that each separable metric space is strongly paracompact, and that a nonseparable metric space need not be strongly paracompact (Example 11). We summarize in Figure 2 these results together with the counterexamples to the converse implications.

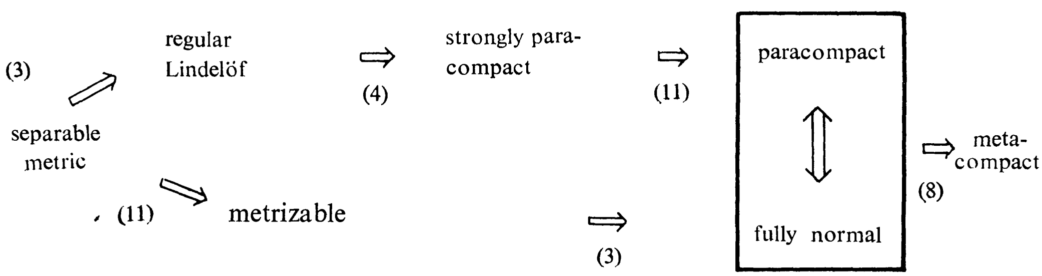


FIG. 2

A most important variation of paracompact spaces is that of **countably paracompact** spaces, those for which every countable open covering has a locally finite

open refinement. Morita **[43]** showed in 1948 that every metacompact normal space is countably paracompact, (see also Michael **[40]**) while in 1951 Dowker **[20]** proved that every perfectly normal space is countably paracompact. Dowker conjectured that every normal space is countably paracompact, and showed this conjecture equivalent to the conjecture that the product of a normal space with the closed unit interval I is normal by showing that X is countably paracompact and normal if and only if $X \times I$ is normal. Countably paracompact normal spaces are sometimes called **binormal**; they have been characterized in many ways by Mansfield **[34]** and Dowker **[20]**. Clearly every fully normal (i.e., paracompact) space is binormal, and every binormal space is normal.

Screenable Spaces. A collection $\mathscr{B}$ of sets is called **conservative** (or **closure preserving**) if for every subcollection $\mathscr{A} \subset \mathscr{B}$, the union of the closure of the members of $\mathscr{A}$ is closed. A conservative collection is **discrete** if the closures are pairwise disjoint. Equivalently a collection $\mathscr{B}$ of subsets of X is discrete if every point in X has a neighborhood which intersects at most one of the sets in $\mathscr{B}$.

Now a topological space is called (by Bing **[10]**) **screenable** if for each open covering F there is a sequence F_n of collections of pairwise disjoint open sets such that $\cup F_n$ is a refinement of F. The space is called **strongly screenable** if the F_n may be chosen to be discrete. A **perfectly screenable** space is one with a **σ-discrete base** — that is, a base which is the countable union of discrete families. A formally weaker condition is that of a **σ-locally finite base** — one which is the countable union of locally finite families. It follows directly from the definitions that every perfectly screenable space is strongly screenable, and *a fortiori*, screenable.

Stone **[53]** showed in 1948 that every metric space has a σ-discrete (and thus σ-locally finite) base. Shortly thereafter, Nagata **[45]** and Smirnov **[50]** showed that every regular space with a σ-locally finite base is metrizable, while Bing **[10]** showed that each perfectly screenable regular space is metrizable. A few years after Bing's work appeared, Nagami **[44]** showed that in regular spaces paracompactness is equivalent to strong screenability and that in binormal (i.e., countably paracompact and normal) spaces, screenable implies strongly screenable. Every strongly screenable developable space must be perfectly screenable since the discrete refinements of the development will form a σ-discrete base **[10]**. Thus every paracompact Moore space is metrizable, for by Nagami's theorem such spaces are strongly screenable and developable. Heath **[25]** showed that every screenable G_δ-space (thus every screenable developable space) is metacompact.

We summarize in Figure 3 the major implications for regular spaces (which are really the only ones of interest *vis-à-vis* metrizability). The relevant counterexamples are classified by the Venn diagram in Figure 4.

Collectionwise Normal Spaces. A (Hausdorff) topological space is called **collectionwise normal** if every discrete collection of sets (or, equivalently, closed sets) can be

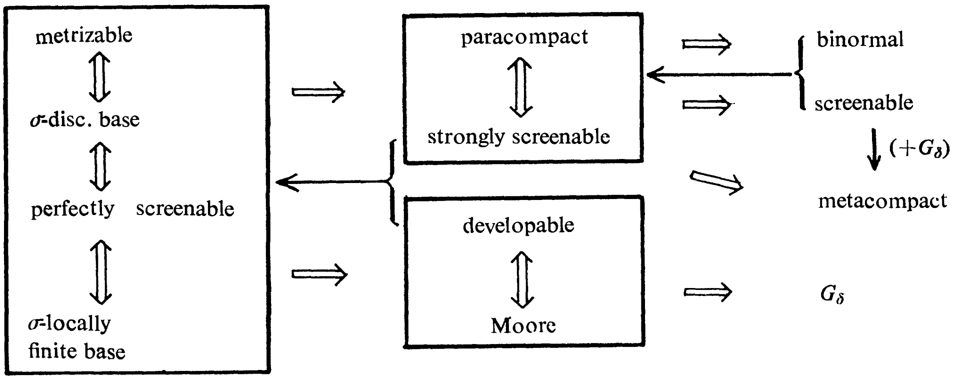


FIG. 3

covered by a pairwise disjoint collection of open sets, each of which covers just one of the original sets. If we weaken this property by requiring it of only countable discrete collections, we call the space **countably collectionwise normal.** On the other hand, we may strengthen collectionwise normal by requiring every **almost discrete** collection of sets (that is, a collection which is discrete with respect to its union)

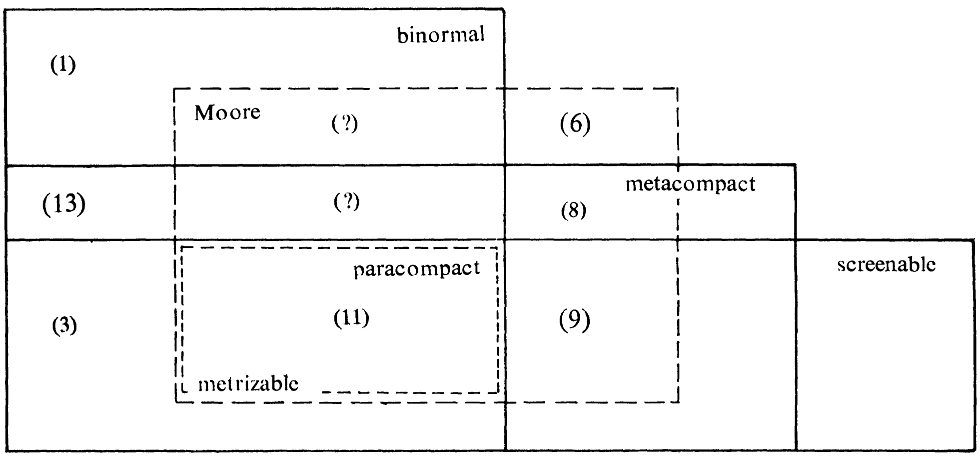


FIG. 4

to have a covering by pairwise disjoint open sets: such spaces are called **completely collectionwise normal**. A space is completely collectionwise normal if and only if it is hereditarily collectionwise normal [**35**], so each completely collectionwise normal space must be completely normal (i.e., hereditarily normal). Every metric space is completely collectionwise normal, so we summarize the implications in Figure 5. Examples 10 and 12 show that normal spaces need not be collectionwise normal, and that collectionwise normal spaces need not be completely collectionwise normal.

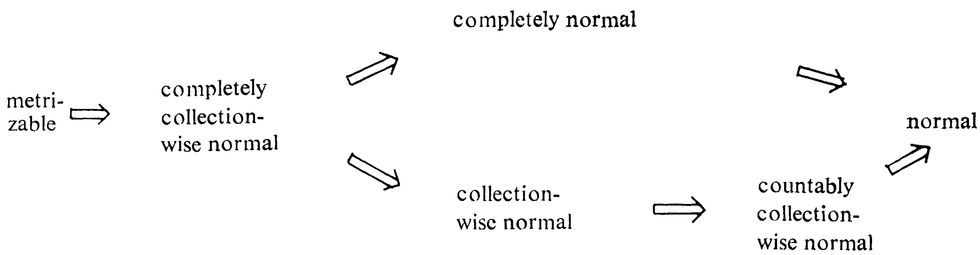


FIG. 5

Bing [**10**] showed that every fully normal (i.e., paracompact) space is collectionwise normal; Nagami [**44**] showed that every metacompact collectionwise normal space is strongly screenable. Nagami and Michael [**38**] showed that the converse holds for regular spaces. So for regular spaces, the concepts of fully normal, paracompact and strongly screenable coincide. Since each strongly screenable developable space is perfectly screenable and each regular perfectly screenable space is metrizable, we conclude again that every paracompact Moore space is metrizable. In fact, Bing [**10**] gave two slightly stronger results: every screenable, normal Moore space is metrizable (since every screenable normal developable space is strongly screenable) and every collectionwise normal Moore space is metrizable (since every such space is screenable). Thus to prove every normal Moore space metrizable, it would suffice to prove it collectionwise normal. In 1964 Bing [**8**] showed that every normal Moore space is countably collectionwise normal.

Several conditional converses of the basic implications have been established. Michael [**40**] showed that every collectionwise normal metacompact space is paracompact, while McAuley [**35**] showed that every collectionwise normal semimetric space is paracompact, and that every paracompact semimetric space is completely collectionwise normal.

In 1960 Alexandroff [**2**] developed a slightly different type of metrization theorem by defining the concept of a uniform base: a basis for X is a **uniform base** if for each $x \in X$ and each neighborhood U of x, only a finite number of the basis sets which contain x intersect $X - U$. Equivalently, a base $\mathscr{B}$ for X is uniform if for each $x \in X$ any infinite subset of $\{U \in \mathscr{B} \mid x \in U\}$ is a (local) basis at x. Since for each integer n the open covering of a metric space by balls of radius $1/n$ has a locally finite subcovering, each metric space has a uniform base, and each space with a uniform base is metacompact. Alexandroff showed that a collectionwise normal space with a uniform base is metrizable, and similarly that a paracompact space with a uniform base is metrizable. Heath [**25**] proved that a regular space has a uniform base if and only if it is metacompact and developable, from which both of Alexandroff's theorems follow.

Arhangel'skii [**5**] strengthened the definition of a uniform base by substituting for the point x an arbitrary compact set K: he called $\mathscr{B}$ a **strongly uniform base** if for

any compact subset $K \subset X$ and any neighborhood U of K, only a finite number of the basis sets intersect both K and $X - U$. Arhangel'skii showed [7] that a space is metrizable if and only if it has a strongly uniform base. Finally, a space is said to have a **point countable base** if it has a basis $\mathscr{B}$ such that no point is contained in more than countably many sets of $\mathscr{B}$. Each uniform base is point countable, and Heath [24] has shown that every semimetric space with a point countable base is developable. We summarize the preceding implications in Figure 6; the reader is invited to draw the corresponding Venn diagram.

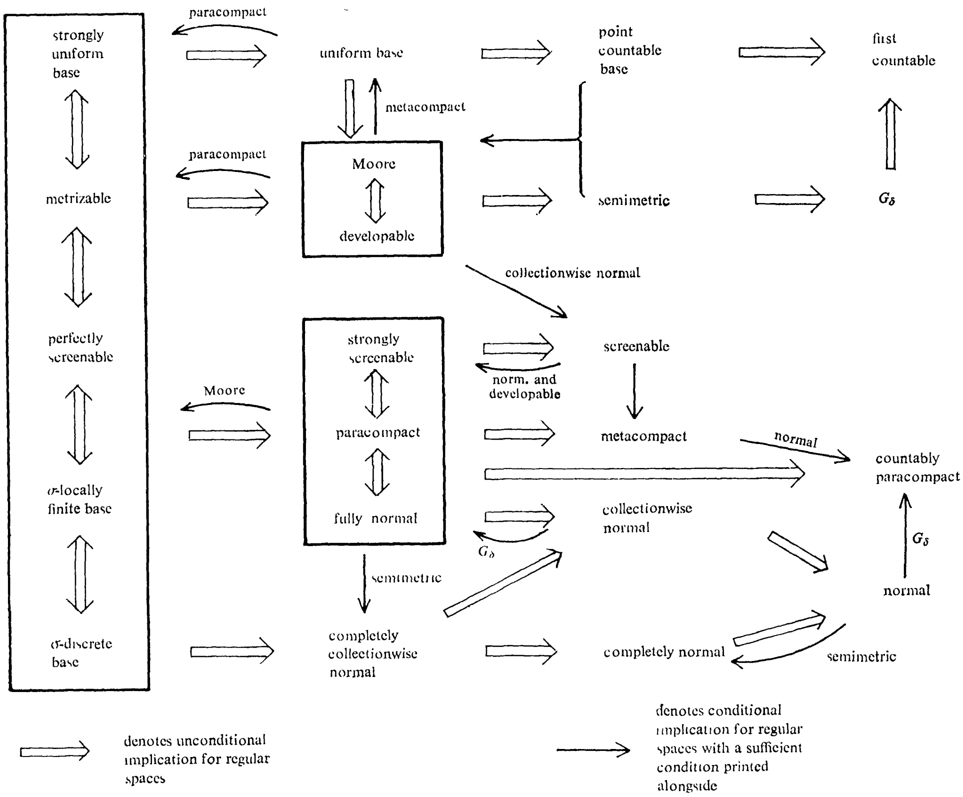


FIG. 6

Conjectures. The literature on the normal Moore space conjecture abounds in conditional theorems which assert that if some hypothesis is true, then some particular theorem is true. A famous example cited previously is Jones' theorem that if $2^{\aleph_0} < 2^{\aleph_1}$, then every separable normal Moore space is metrizable. These theorems deal with implications among statements whose truth or falsehood is either not yet known, or which are in some cases (e.g., the continuum hypothesis) independent of the axioms of set theory.

We shall denote by CH the continuum hypothesis $2^{\aleph_0} = \aleph_1$; Gödel [22] and Cohen

[18] proved this hypothesis consistent with and independent of the Zermelo-Fraenkel (or Gödel-Bernays) axioms of set theory (hereafter referred to simply as "set theory"). We shall denote by ***WCH*** Jones' hypothesis that $2^{\aleph_0} < 2^{\aleph_1}$, since it is a weak version of CH: if $2^{\aleph_0} = \aleph_1$, then $2^{\aleph_0} = \aleph_1 < 2^{\aleph_1}$ by Cantor's theorem. Clearly the consistency of CH implies the consistency of WCH. The negation of WCH, namely $2^{\aleph_0} = 2^{\aleph_1}$ is called the Luzin Hypothesis (***LH***); Bukovsky [14] showed that LH is consistent with set theory. Thus WCH, the negation of LH, is independent of set theory.

Since every separable metric space has $2^{\aleph_0}$ Borel subsets WCH implies that every separable uncountable metric space has a subset which is not a Borel set; we shall call this ***BH***, for Borel hypothesis. Heath [25] used a special case of BH to strengthen Jones' theorem: we shall denote by ***HH*** the statement that every uncountable subspace M of the real line contains a subset which is not F_σ in M. Since every F_σ-set is a Borel set, BH implies HH; Heath showed that HH is equivalent to Jones' conjecture ***JC*** that every separable normal Moore space is metrizable. The consistency of the continuum hypothesis implies that of JC, while the independence of JC was proved by Tall and Silver [54] in 1970.

Heath also showed that Jones' conjecture follows from the hypothesis ***MMSC*** that every normal metacompact Moore space is metrizable; clearly $MMSC$ is weaker than the normal Moore space conjecture ***MSC***. $MMSC$ is equivalent to Alexandroff's conjecture ***AC*** that every normal space with a uniform base is metrizable [3]. Traylor [57] suggested the conjecture (***TC***) that every normal Moore space is metacompact. Since McAuley [35] showed that a separable normal metacompact Moore space is metrizable, Traylor's conjecture implies Jones' conjecture.

Several common conjectures center on semimetric spaces, a generalization of Moore spaces. Brown [13] suggested that every normal semimetric space is collectionwise normal, while Heath [23] appeared to strengthen this conjecture by suggesting that every normal semimetric space is paracompact. Actually since every semimetric collectionwise normal space is paracompact [35], these conjectures are equivalent; we shall denote them by ***NSP***. McAuley [37] proposed the weaker conjecture $SNSP$ that every separable normal semimetric space is paracompact. The Bing-Nagami result that every paracompact Moore space is metrizable shows that NSP implies the Moore space conjecture MSC, and similarly, $SNSP$ implies Jones' separable Moore space conjecture JC.

In [10] Bing showed that MSC is equivalent to the conjecture that every normal Moore space is collectionwise normal; in [8], he considered the weaker conjecture ***BC*** that every normal Moore space is collectionwise normal with respect to a discrete collection of points. (He termed a counterexample to BC one of *type D*.) Bing showed that BC is equivalent to the following set theoretic conjecture: If X is a set and if Y denotes the product $X \times X$ less the diagonal $\Delta = \{(x,x) \in X \times X\}$, we call a subset $W \subset Y$ a **skew subset** if the projections $\pi_x(W)$ and $\pi_y(W)$ are disjoint. Bing's alternative to BC is the conjecture ***F*** that if $f: Y \to Z^+$ is a function from Y to the nonnegative integers with the property that for each skew subset $W \subset Y$ there is a function

$F_W: W \to Z^+$ which dominates f in the sense that $\max\,[F_W(x), F_W(y)] > f(x,y)$ for all $(x,y) \in W$, then there is a function $F: X \to Z^+$ which dominates f in this sense for all $(x,y) \in Y$.

Bing also showed that BC implies JC by showing that any nonmetrizable separable normal Moore space would necessarily be a counterexample of type D. We summarize the relationships among these conjectures in Figure 7. Since all of the conjectures in this figure imply JC, none of them can be proved from the axioms of set theory. But the consistency of these various hypotheses (except of course for CH and its consequences) remains an open question.

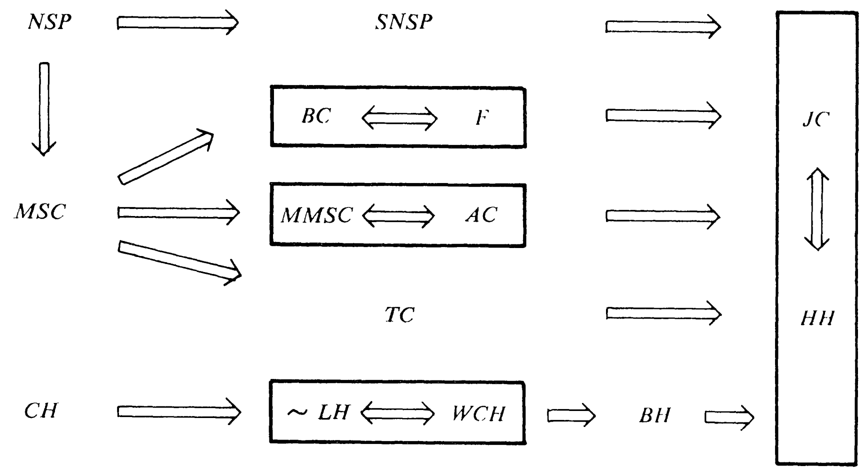


FIG. 7

We have already mentioned Dowker's conjecture ***DC*** that every normal space is countable paracompact; he showed this equivalent to the conjecture ***NP*** that the product of every normal space with the unit interval is normal **[18]**. Nagami **[44]** showed that a screenable normal countably paracompact space is paracompact and conjectured ***NC*** that every screenable normal space is paracompact. Clearly DC implies NC.

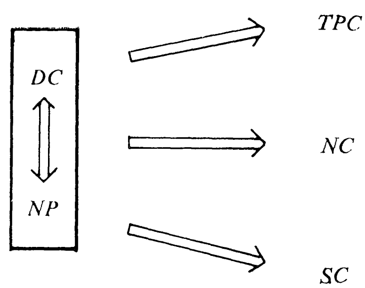


FIG. 8

Tamano [**55**] discusses a wide variety of theorems concerning the product invariance of normality and paracompactness and enunciates the following conjecture ***TPC***: If Y is metrizable and $X \times Y$ is normal then $X \times Y$ is paracompact. Tamano and Morita [**42**] have shown that to conclude that $X \times Y$ is paracompact it is sufficient to prove $X \times Y$ countably paracompact . Thus Dowker's conjecture implies Tamano's.

Souslin [**50**] asked whether a linearly ordered space must be separable whenever it satisfies the countable chain condition (that every disjoint collection of open sets is at most countable). We shall call this conjecture ***SC***; a counterexample (if it exists) is known as a **Souslin space**. A thorough discussion of this conjecture and related topics is provided by M. E. Rudin [**47**] who earlier showed [**46**] that if a Souslin space exists, then so must a counterexample to Dowker's conjecture. In other words, Dowker's conjecture implies Souslin's conjecture. Tennenbaum, Solovay, and Jech showed that Souslin's conjecture is consistent with [**49**] and independent of ([**27**, **56**]) the axioms of set theory. Thus Dowker's conjecture cannot be proved from the present axioms of set theory. (*Added in proof:* In fact, it is false. Just recently M. E. Rudin constructed a counterexample to Dowker's conjecture.)

Epilogue. The concepts and examples discussed in this paper represent not so much the frontier as the established settlements of metrization research. Several recent papers by Ceder [**16**], Borges [**11**], [**12**], Michael [**39**], and Worrell and Wicke [**62**] contain such refinements as M_i-spaces, stratifiable spaces, $\aleph_0$-spaces, and θ bases. In each of these new areas there are significant and difficult conjectures similar to those enumerated above; the interested reader can pursue these issues in the papers cited in the bibliography, together with those listed in the excellent bibliographies of [**3**] and [**6**].

Since a metric is a map to the positive reals, it should not be surprising to find that the existence of certain esoteric metrics is intimately related to the existence of certain subsets of the real line. Example 7 provides a very specific instance of this relationship in that potential counterexamples to both Jones' and Dowker's conjectures depend on the existence of certain special subsets of the real line, while the independence theorems of Tall, Silver, Tennenbaum, Solovay, and Jech show that many topological problems depend on fundamentally undecidable problems of set theory. Thus many of the unresolved metrization conjectures may come to be viewed as one measure of the incompleteness of our present axiomatic view of metric spaces.

Examples

1. *Open Ordinal Space.* Let X be the set of all ordinal numbers strictly less than the first uncountable ordinal Ω; X carries the interval (or order) topology. Then X is completely collectionwise normal [**51**] but not fully normal [**10**].

2. *Closed Ordinal Space.* Let X be the set of all ordinal numbers less than or equal to the first uncountable ordinal Ω. X is compact in the interval topology, but not G_δ

since the closed set $\{\Omega\}$ is not a G_δ set. Thus X is neither perfectly normal nor semimetrizable. But of course it is strongly paracompact.

3. *Lower Limit Topology*. Let X be the real line with the topology generated by the sets of the form $[a, b) = \{x \in X \mid a \leqq x < b\}$. Bing [**10**] cites this space as an example of a regular, separable, strongly screenable (and therefore paracompact) space which is neither perfectly screenable nor developable.

4. *Stratified Plane*. If R is the real line with the Euclidean topology and S is the real line with the discrete topology, then $X = R \times S$ is a nonseparable strongly paracompact metric space.

5. *Bow-Tie Space*. Let X be the Euclidean plane with real axis L. If $d: X \times X \to R^+$ is the Euclidean metric on X, we define a semimetric δ as follows: $\delta(p,q) = d(p,q)$ if $p,q \in X - L$; $\delta(p,q) = d(p,q) + \alpha(p,q)$ if p or $q \in L$, where $\alpha\ (p,q)$ is the radian measure of the acute angle between L and the line connecting p to q. The topology on X is generated by the semimetric balls of small radius; a neighborhood ball of a point $p \in L$ looks like a bow-tie (Figure 9) or a butterfly, so this space is often called

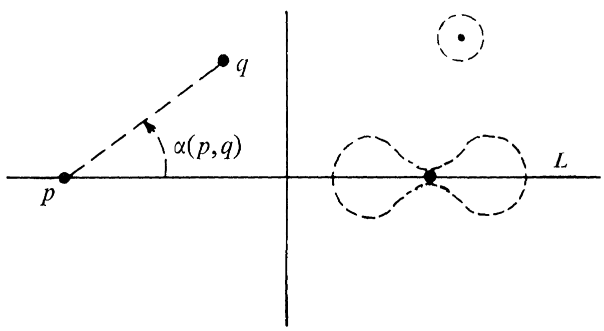


FIG. 9

the bow-tie or butterfly space. McAuley [**36**] introduced this space as an example of a regular semimetric space which is not developable. He showed furthermore that it is paracompact (thus completely collectionwise normal) and hereditarily separable.

6. *Tangent Disc Topology*. Let $P = \{(x, y) \mid x, y \in R, y > 0\}$ be the open upper half-plane with the Euclidean topology τ and let L denote the real axis. We generate a topology on $X = P \cup L$ by adding to τ all sets of the form $\{x\} \cup D$, where $x \in L$ and D is an open disc in P which is tangent to L at the point x (Figure 10). This important example was apparently introduced by both Niemytzki (see [**6**]) and Moore (see [**29**]) as a regular developable space which is not metrizable (since the uncountable closed subset L is discrete and thus not separable in the induced topology). The development which makes X a Moore space is the collection of

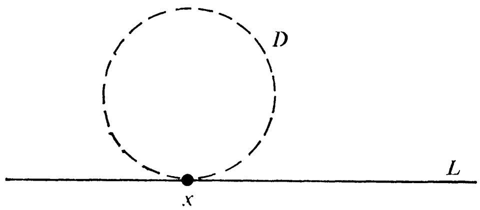


FIG. 10

open balls of radius $1/n$ (including the tangent discs $\{x\} \cup D$ if D has radius $1/n$). X is clearly not normal, and neither countably paracompact nor metacompact [52].

A common variation (see [30]) of the tangent disc topology is formed by replacing the tangent disc neighborhoods by sets of the form $\{x\} \cup T$ for each $x \in L$, where T is an inverted isosceles triangle in P with vertex at the point x and base parallel to L, such that the radian measure of the vertex angle equals the length

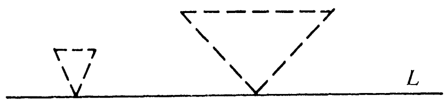


FIG. 11

of its adjacent sides (Figure 11). McAuley [36] discusses a different variation which is formed from the bow-tie space by rotating each of the bow-tie neighborhoods 90° (Figure 12). Bing [9] introduced a physical model which he called flow space by assuming that water is flowing from left to right across the unit square at the rate of $(1-x)$ feet per second. Flow space is the closed unit square, and a neighborhood $N_p(t)$ of a point p is the set of all points in X which a swimmer could reach in less than t seconds (Figure 13).

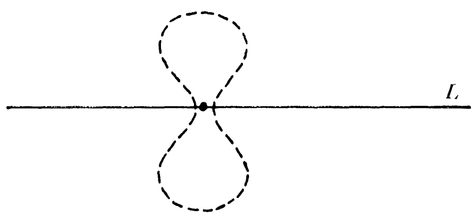


FIG. 12

7. *Tangent Disc Subspaces*. If S is a subset of the real line L, and $Y = P \cup L$ is the tangent disc space, we let X be the subspace $P \cup S$ with the topology induced from Y. The space X is second countable if and only if S is countable, so, since X is regular, X is metrizable if and only if S is countable. X will always be a Moore space since it has the same development as Y, and similarly it will always be separable since the rational lattice points in P are dense in X. Jones [28] showed that every subset of cardinality c of a separable normal space has a limit point; since S cannot

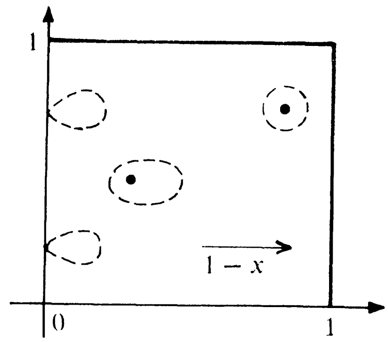


FIG. 13

have any limit points, X cannot be normal when S has cardinality c. Bing [**10**] showed that X is normal if every subset of S is a G_δ-set in (the relative topology of) S; but every uncountable G_δ-subset of the Euclidean real line has cardinality c (by Mazurkiewicz' theorem [**33**, p. 441]). Thus X would be a normal nonmetrizable Moore space if S were uncountable but of cardinality less than c with the additional property that every subset of S is G_δ in S. Such an S could contain only countable G_δ-subsets of the real line. Clearly the existence of a set with these properties cannot be proved within ordinary set theory since it would constitute a counterexample to the continuum hypothesis. However, Jones [**39**] constructed a set S of cardinality $\aleph_1$ such that every countable subset of S is G_δ in S.

Younglove [**63**] studied this example as a possible counterexample to Dowker's conjecture that every countably paracompact space is normal and proved that if S is a G_δ-set, then X is countably paracompact if and only if S is countable. Thus X could be a counterexample to Dowker's conjecture only if S was not a G_δ-subset of the real line L.

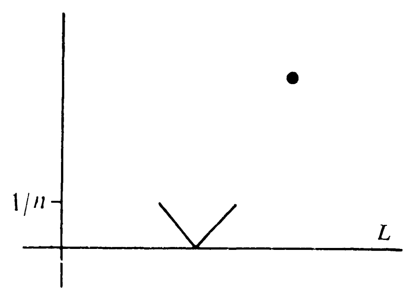


FIG. 14

8. *Tangent V Topology*. If X is the upper half plane including the real axis L, we let each point of $X - L$ be open and take as a neighborhood basis of points $x \in L$ a "V" with vertex at x, sides of slopes ± 1 and height $1/n$ (Figure 14). Heath [**25**]

showed that X is a metacompact Moore space which is not screenable. Clearly X is neither normal nor separable.

9. *Picket Fence Topology*. If X is the upper half plane including the real axis L, we let each point of $X - L$ be open, and take as a neighborhood basis of rational points $x \in L$ the vertical line segments of height $1/n$ with lower end point at x. The neighborhood basis of irrational points $x \in L$ consists of line segments of slope 1 and height $1/n$ with their base at the point x (Figure 15). Heath [**25**] introduced this as a simple example of a screenable Moore space which is not normal.

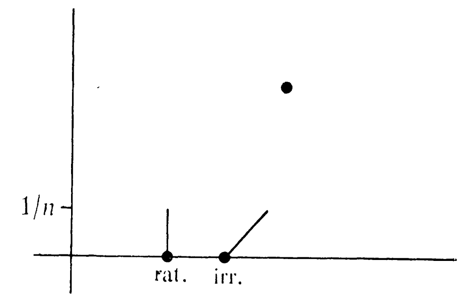


FIG. 15

10. I^I. Let $X = I^I$ be the uncountable Cartesian product of the closed unit interval $I = [0,1]$ with the Tychonoff topology; that is, X is the set of all functions from I to I with the topology of pointwise convergence. Since X is compact and Hausdorff, it is normal; but it is not completely normal [**52**] since it contains a subspace homeomorphic to Z^I, the uncountable product of the positive integers, which Stone [**53**] showed was not normal. Thus X is strongly paracompact and collectionwise normal but neither perfectly normal nor developable.

11. *Hedgehog*. If K is a cardinal number, a hedgehog X of spininess K is formed from the union of K disjoint copies of the unit interval [0,1] by identifying the zero points of each interval. A metric for X can be defined by $d(x, y) = |x - y|$ if x and y belong to the same segment (or spine), and $d(x,y) = x + y$ otherwise. Alexandroff [3] cites a hedgehog of uncountable spininess as an example of a metric space which is not strongly paracompact.

12. *Bing's Power Space*. If S is some uncountable set with power set P, let $X = \prod_{\lambda \in P}\{0,1\}_\lambda$, where $\{0,1\}_\lambda$ is a copy of the two point discrete space. (If we let 2 denote the two point discrete space, we have $X = 2^{2^S}$.) Since the elements of X are collections of subsets of S, each ultrafilter on S is a point in X; let M denote the subset of X consisting of all principal ultrafilters of S. Then if x_s is the point in X whose λ-th coordinate $(x_s)_\lambda$ equals 1 if and only if $s \in \lambda$, we have $M = \{x_s \in X \mid s \in S\}$. If X has the

Tychonoff topology τ, $X - M$ is dense in X. Bing **[10]** generated a new topology on X by adding to τ all points of $X - M$ as open sets; we shall denote the topology thus generated by σ. M inherits from (X,σ) the discrete topology; furthermore, any two disjoint closed subsets of M are contained in disjoint open subsets of X **[52]**. It follows that X is normal but not perfectly normal **[10]**, metacompact **[52]** or collectionwise normal (since M is an uncountable discrete collection of points without disjoint open neighborhoods for all of its points).

13. *Michael's Power Subspace.* If $X = 2^{2^S}$ is Bing's Power Space, we let Y be the subspace $M \cup L$, where M is the subset of all principal ultrafilters of S and L is the collection of all finite families in $X-M$. Michael **[40]** selected this subspace as an example of a normal metacompact space which is not collectionwise normal.

14. *Cantor Tree.* Let C denote the Cantor set in the unit interval $[0,1]$; the midpoints of the components of $[0,1]-C$ are $1/2$, $1/6$, $5/6$, $1/18$, $5/18$, etc. Let D be the tree (or dendron) in the lower half plane whose vertices are $(1/2,-1)$, $(1/6,-1/2)$, $(5/6,-1/2)$, $(1/18,-1/4)$, $(5/18,-1/4)$, etc.

Then the space X is defined as $D \cup C$ (Figure 16), where D inherits the Euclidean topology from the plane, while a basis neighborhood of a point $c \in C$ is a path Γ in the tree D whose upper limit is the point c, together with open segments at each branch point of Γ sufficiently short to avoid including any other branch point. Jones **[31]** cites this example of Moore as the first example of a nonmetrizable Moore space. The fact that X is nonmetrizable follows from the observation that it is separable but not perfectly separable. Jones **[31]** shows that X is not normal.

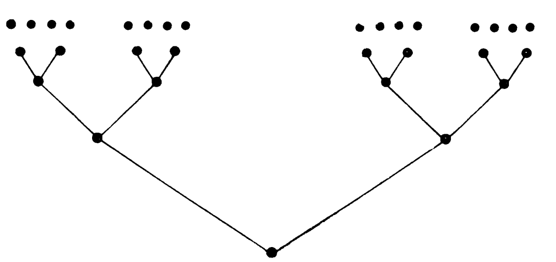


FIG. 16

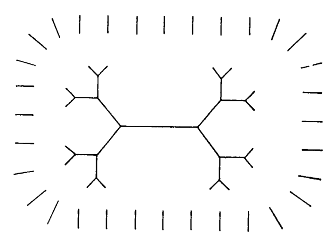


FIG. 17

15. *Moore's Road Space.* Let two roads start at the origin of the plane and proceed in opposite directions for one mile each. Let each then branch into two roads which continue for one mile each before each of these now branches into two roads. Continue this in such a way that none of the new roads ever intersect, and so that all roads proceed indefinitely far from the origin. This process generates c roads; at the "end" of each we adjoin a straight ray of infinite length. This collection of roads is the space X (Figure 17), and we generate a topology from a basis of open discs. This

"automobile road" space was introduced by Moore as a graphic variation of the Cantor tree (Example 14); it has the same properties [31].

1 = *example has property*

0 = *example does not have property*

Example	Hausdorff	Regular	Normal	Comp. Normal	Perf. Normal	Coll. Normal	Count. Col. Nor.	Comp. Col. Nor.	Metrizable	Separable	2nd Count.	1st Count.	Hered. Separ.	Lindelöf	Developable	Moore	Semimetric	G_δ	Fully Normal	Str. Paracompact	Paracompact	Metacompact	Count. Paracmpt.	Binormal	Screenable	Str. Screenable	Perf. Screenable	σ-loc. Fin. Base	Uniform Base	Pt. Count. Base
1 Open Ord. Space	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	
2 Closed Ord. Sp.	1	1	1	1	0	1	1	1	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	
3 Lower Limit	1	1	1	1	1	1	1		0	1	0	1		1	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	
4 Stratified Plane	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5 Bow-Tie	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	
6 Tangent Disc	1	1	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	
8 Tangent V	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	1		0	0	0	0	0	1	1
9 Picket Fence	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	1		0	1	0	0	0	1	1
10 I^I	1	1	1	0	0	1	1	0	0	1	0	0		1	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	
11 Hedgehog	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
12 Power Space	1	1	1		0	0		0	0		0			0	0	0	0	0	0	0	0	0				0	0	0	0	
13 Power Subspace	1	1	1		0	0		0	0		0					0	0	0	0	0	0	1	1	1	0	0	0	0		
14 Cantor Tree	1	1	0	0	0	0	0	0	0	1	0	1		0	1	1	1	1	0	0	0				0	0	0	0		
15 Moore's Road Space	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0				0	0	0	0		

FIG. 18

References

1. P. Alexandroff, General duality theorems for non-closed sets of n-dimensional space, Mat. Sbornik, 21 (63) (1947) 161–231.

2. ———, On the metrization of topological spaces, Bull. Acad. Polon. Sci., 8 (1960) 135–140.

3. ———, Some results in the theory of topological spaces, obtained within the last twenty-five years, Russian Math. Surveys, 15 (1960) 23–83.

4. P. Alexandroff and P. Urysohn, Une condition nécessaire et suffisante pour qu'une classe (L) soit une classe (B), Comptes Rendus, 177 (1923) 1274–1276.

5. A. Arhangel'skii, Bicompact sets and the topology of spaces, Soviet Math. Dokl., 4 (1963), 561–564.

6. ———, Mappings and spaces, Russian Math. Surveys, 21 (1966) 115–162.

7. ———, On the metrization of topological spaces, Bull. Acad. Polon. Sci., 8(1960) 589–595.

8. R. H. Bing, A translation of the normal Moore space conjecture, Proc. Amer. Math. Soc., 16 (1965) 612–619.

9. ———, Challenging conjectures, this MONTHLY, 50th Anniv. Issue, 74 (1967) 56–64.

10. ———, Metrization of topological spaces, Canad. J. Math., 3 (1951) 175–186.

11. C. R. Borges, On the metrizability of topological spaces, Canad. J. Math., 20 (1968) 795–804.

12. ———, Stratifiable spaces, Pacific J. Math., 17 (1966) 1–16.

13. M. Brown, Semi-metric spaces, Summer Institute on Set Theoretic Topology, Madison, A.M.S., (1955) 64–66.

14. L. Bukovsky, Borel subsets of metric separable spaces, General Topology and its Relations to Modern Analysis and Algebra, 1966 Prague Symp., Academic Press, 1967, 83–86.

15. E. Čech, Sur la dimension des espaces parfaitement normaux, Bull. Intern. de l'Acad. de Bohême (Prague), 33 (1932) 38–55.

16. J. G. Ceder, Some generalizations of metric spaces, Pacific J. Math., 11 (1961) 105–125.

17. E. W. Chittenden, On the metrization problem and related problems in the theory of abstract sets, Bull. Amer. Math. Soc., 33 (1927) 13–34.

18. P. J. Cohen, The independence of the continuum hypothesis, I, II, Proc. Nat. Acad. Sci. U.S.A., 50 (1963) 1143–1148; 51 (1964) 105–110.

19. J. Dieudonné, Une généralisation des espaces compacts, J. Math. Pures Appl., 23 (1944) 65–76.

20. C. H. Dowker, On countably paracompact spaces, Canad. J. Math., 3 (1951) 219–224.

21. S. Gaal, Point Set Topology, Academic Press, New York, 1964.

22. K. Gödel, The Consistency of the Continuum Hypothesis, Princeton, 1940.

23. R. W. Heath, On certain first countable spaces, Topology Seminar, Wisconsin (1965); Princeton (1966) 103–113.

24. ———, On spaces with point-countable bases, Bull. Acad. Polon. Sci., 13 (1965) 393–395.

25. ———, Screenability, pointwise paracompactness, and metrization of Moore spaces, Canad. J. Math., 16(1964) 763–770.

26. ———, Separability and $\aleph_1$ compactness, Colloq. Math., 12 (1964) 11–14.

27. T. Jech, Non-provability of Souslin's Hypothesis, Comm. Math. Univer. Carolinae, 8 (1967) 291–305.

28. F. B. Jones, Concerning normal and completely normal spaces, Bull. Amer. Math. Soc., 43 (1937) 671–677.

29. ———, Metrization, this MONTHLY, 73 (1966) 571–576.

30. ———, Moore spaces and uniform spaces, Proc. Amer. Math. Soc., 9 (1958) 483–486.

31. ———, Remarks on the normal Moore space metrization problem, Topology Seminar, Wisconsin (1965); Princeton (1966) 115–119.

32. S. Kaplan, Homology properties of arbitrary subsets of Euclidean spaces, Trans. Amer. Math. Soc., 62 (1947) 248–271.

33. C. Kuratowski, Topologie I, Monografie Matematyczne, Vol. 20, Warsaw, 1958.

34. M. J. Mansfield, On countably paracompact normal spaces, Canad. J. Math., 9 (1957) 443–449.

35. L. F. McAuley, A note on complete collectionwise normality and paracompactness, Proc. Amer. Math. Soc., 9 (1958) 796–799.

36. ———, A relation between perfect separability, completeness, and normality in semi-metric spaces, Pacific J. Math., 6 (1956) 315–326.

37. ———, Paracompactness and an example due to F. B. Jones, Proc. Amer. Math. Soc., 7 (1956) 1155–1156.

38. E. A. Michael, A note on paracompact spaces, Proc. Amer. Math. Soc., 4 (1953) 831–838.

39. ———, $\aleph_0$ spaces, J. Math. Mech., 15 (1966) 983–1002.

40. ———, Point finite and locally finite coverings, Canad. J. Math., 7 (1955) 275–279.

41. R. L. Moore, Foundations of point set theory, Amer. Math. Soc. Coll. Publ. 13, New York, 1932.

42. K. Morita, Products of normal spaces with metric spaces, Math. Ann., 154 (1964) 365–382.

43. ———, Star-finite coverings and the star-finite property, Math. Japon., 1 (1948) 60–68.

44. K. Nagami, Paracompactness and strong screenability, Nagoya Math. J., 8(1955) 83–88.

45. J. Nagata, On a necessary and sufficient condition of metrizability, J. Inst. Polytech., Osaka City Univ., 1 (1950) 93–100.

46. M. E. Rudin, Countable paracompactness and Souslin's problem, Canad. J. Math., 7 (1955) 543–547.

47. ———, Souslin's conjecture, this Monthly, 76 (1969) 1113–1119.

48. Yu. M. Smirnov, On metrization of topological spaces, Uspehi Mat. Nauk., 6 (1951) 100–111 (A.M.S. Transl. No. 91).

49. R. Solovay and S. Tennenbaum, Iterated Cohen extensions and Souslin's problem, Fund. Math.

50. M. Souslin, Problème 3, Fund. Math., 1 (1920) 223.

51. L. Steen, A direct proof that a linearly ordered space is hereditarily collectionwise normal, Proc. Amer. Math. Soc., 24 (1970) 727–728.

52. L. Steen and J. A. Seebach, Counterexamples in Topology, Holt, Rinehart and Winston, New York, 1970.

53. A. H. Stone, Paracompactness and product spaces, Bull. Amer. Math. Soc., 54 (1948) 977–982.

54. F. D. Tall, New results on the normal Moore space problem, Proc. of Washington State Univ. Conf. on General Topology, 1970, 120–126.

55. H. Tamano, Normality and product spaces, General Topology and its Relations to Modern Analysis and Algebra, 1966 Prague Symp., Academic Press, 1967, 349–352.

56. S. Tennenbaum, Souslin's Problem, Proc. Nat. Acad. Sci. U. S. A., 59 (1968) 60–63.

57. D. R. Traylor, On normality, pointwise paracompactness and the metrization question, Topology Conference, Arizona State Univ., (1967) 286–292.

58. J. W. Tukey, Convergence and uniformity in topology, Princeton, 1940.

59. A. Tychonoff, Über einen Metrisationssatz von P. Urysohn, Math Ann., 95 (1926) 139–142.

60. P. Urysohn, Zum Metrisationsproblem, Math. Ann., 94 (1925) 309–315.

61. C. W. Vickery, Axioms for Moore spaces and metric spaces, Bull. Amer. Math. Soc., 46 (1940) 560–564.

62. J. M. Worrell, Jr., and H. H. Wicke, Characterizations of developable topological spaces, Canad. J. Math., 17 (1965) 820–830.

63. J. N. Younglove, Two conjectures in point set theory, *Topology Seminar*, Wisconsin (1965), Princeton (1966), 121–123.

THE ORIGINS OF MODERN AXIOMATICS: PASCH TO PEANO

H. C. KENNEDY, Providence College

The modern attitude toward the undefined terms of an axiomatic mathematical system is that popularized by Hilbert's remark: "One must be able to say at all times —instead of points, straight lines, and planes—tables, chairs, and beer mugs." [**20**, p. 57.] This view was not widely accepted before the twentieth century, and even in 1959 the well-known James and James *Mathematics Dictionary* gave "A self-evident and generally accepted principle" as first meaning of the term "axiom," although this may only be meant as a reflection of the view universally accepted before the developments in geometry in the nineteenth century. The change in attitude appears to be due to internal pressures within mathematics (what R. L. Wilder has called "hereditary stress" [**22**, p. 170]). These include the flowering of projective geometry and, especially, the discovery of the non-euclidean geometries, i.e., of the possibility of a geometry based on axioms, one of which is the negation of one of Euclid's axioms. The transition from viewing an axiom as "a self-evident and generally accepted principle" to the modern view took place in the second half of the nineteenth century and can be found in the very brief period from 1882 to 1889, from Pasch's *Vorlesungen über neuere Geometrie* [**13**], to Peano's *I principii di Geometria logicamente esposti*, [**15**].

Already in 1882, Pasch showed a shift in interest from the theorems to the axioms from which the theorems are derived, when he insisted that everything necessary to deduce the theorems must be found among the axioms [**13**, p. 5]. Pasch was concerned that his axiom set be complete, i.e., that it furnish a basis for rigorous proofs of the theorems. ("The father of rigor in geometry is Pasch", wrote Hans Freudenthal [**5**, p. 619].) There is also a strong hint of the modern attitude, as expressed in Hilbert's remark about "tables, chairs, beer mugs" in his statement: "In fact, provided the geometry is to be truly deductive, the process of inference must be entirely independent of the *meaning* of the geometrical terms, just as it must be independent of the figures" [**13**, p. 98].

Hilbert's remark was made to a few friends in the waiting room of a railway station in 1891 but was not published until 1935 [**10**, p. 403]. The exposition of the axioms in his famous *Grundlagen der Geometrie* [7] begins: "Let us consider three distinct systems of things. The things composing the first system, we will call *points*, and designate them by the letters $A, B, C, \ldots$" [**8**, p. 3]. The viewpoint is quite clear—but he was not the first to publish this view. Pasch has already been mentioned, and Hans Freudenthal, in a study of geometrical trends at the turn of the century says: "Hilbert had in this view too at least one forerunner, namely G. Fano, . . . " [**4**, p. 14]. He refers to Fano's statement: "As basis for our study we assume an arbitrary *collection* of entities of an arbitrary nature; entities which, for brevity, we shall call *points*, and this quite independently of their nature." [**3**, p. 108.]

Somewhat suprisingly Freudenthal overlooks Peano's monograph of 1889, even though it is citedin Fano's article, perhaps because Fano says that Peano's work was based on that of Pasch. Peano's work was indeed based on his reading of Pasch, but there are important innovations, and one of them is the explicit statement of the modern attitude toward the undefined terms of an axiomatic mathematical system. The first line of his exposition is: "The sign **1** is read **point,**" and in his commentary he says: "We thus have a category of entities, called points. These entities are not defined. Also, given three points, we consider a relation among them, indicated by $c \varepsilon ab$, and this relation is likewise undefined. The reader may understand by the sign **1** any category whatever of entities, and by $c \varepsilon ab$ any relation whatever among three entities of that category, . . . " [**15**; **18**, p. 77]. We find in this statement explicit acceptance of the axiomatic view. (It should be noted that Peano's view was purely methodical. As we have indicated elsewhere [**12**, p. 264], he was not a member of what came to be called the 'formalist' school.)

E. W. Beth has noted [**2**, p. 82]: "Since the publication of D. Hilbert's *Grundlagen der Geometrie* (1899), it has become customary to require every set of axioms to be (1) *complete*, (2) *independent*, and (3) *consistent*." Again, it was Hilbert who popularized this 'custom', but that these properties of an axiom set are desirable was already accepted by Peano and others. The property of consistency is indeed a *sine qua non*, but as the consistency of Euclid's axioms was never doubted, it was only with advent of non-euclidean geometry that attention was focused on this property, and it was not until 1868 that a consistency proof was found by E. Beltrami [**1**]. The property of independence can be reduced to that of consistency; we often say that Beltrami proved the independence of Euclid's "parallel postulate", but this reflects a later view, that of Peano who developed this technique into a general method.

Peano's acceptance of the goal of an independent set of axioms is indicated in his *I principii di Geometria:* "This ordering of the propositions clearly shows the value of the axioms, and we are morally certain of their independence" [**15**; **18**, p. 57]. In a similar remark about his axioms for the natural numbers, published earlier that year, Peano later wrote: "I had moral proof of the independence of the primitive propositions from which I started, in their substantial coincidence with the definitions of Dedekind" [**17**; **19**, p. 243]. It was only in 1891, however, after he had separated the 'famous five' from the postulates dealing with the symbol =, that he showed their absolute independence [**16**; **19**, p. 87].

Hermann Weyl wrote of Hilbert [**20**, p. 264]: "It is one thing to build up geometry on sure foundations, another to inquire into the logical structure of the edifice thus erected. If I am not mistaken, Hilbert is the first who moves freely on this higher 'metageometric' level: systematically he studies the mutual independence of his axioms and settles the question of independence from certain limited groups of axioms for some of the most fundamental geometric theorems. His method is the *construction of models:* the model is shown to disagree with one and to satisfy all other axioms; hence the one cannot be a consequence of the others." This method was, as we have

seen, already used systematically by Peano, although one would not learn this from reading Hilbert. In the *Grundlagen der Geometrie* there is no mention of Peano. The only Italian mentioned is G. Veronese, and the reference is to a German translation of his work. Nor does Hilbert mention Peano even in his presentation of postulates for the real numbers [**9**]. Indeed (without naming him) he labels Peano's development of the real numbers the "genetic method," while reserving the label "axiomatic method" for his own presentation!

A word more may be said about the originality of Peano's work. In contrast with Hilbert, Peano always tried to place his work in the historical evolution of mathematics, to see it as a continuation and development of the work of others. Furthermore he was scrupulously honest (although sometimes mistaken) in assigning priority of discovery. Thus in *I principii di Geometria* he praises Pasch's book and indicates precisely to what extent his treatment coincides with that of Pasch, and where it differs. On the other hand, Peano's discovery of the postulates for the natural numbers was entirely independent of the work of Dedekind, contrary to what is often supposed. Jean van Heijenoort says [**6**, p. 83]: "Peano acknowledges that his axioms come from Dedekind," referring the reader to the statement of Peano: "The preceding primitive propositions are due to Dedekind." [**16, 19**, p. 86]. Hao Wang says [**21**, p. 145]: "It is rather well known, through Peano's own acknowledgement... that Peano borrowed his axioms from Dedekind. . . .," and he gives a reference to Jourdain [**11**, p. 273], which in turn refers to the same passage of Peano just quoted. Since Peano had already written in *Arithmetices Principia:* "Also quite useful to me was a recent work: R. Dedekind, *Was sind und was sollen die Zahlen*, Braunschweig, 1888" [**14; 18**, p. 22], the conclusion of these authors would seem justified. In fact, Peano was only acknowledging Dedekind's priority of publication.

The exact story was given in 1898 when Peano wrote: "The composition of my work of 1889 was still independent of the publication of Dedekind just mentioned; before it was printed I had moral proof of the independence of the primitive propositions from which I started, in their substantial coincidence with the definitions of Dedekind. Later I succeeded in proving their independence," [**17; 19**, p. 243]. We see from this that the reference to Dedekind's work was added to the preface of *Arithmetices Principia* just before the pamphlet went to press, and we have an explanation of how Dedekind's work was "useful".

Ironically, the very modesty of Peano and his desire to see his work as in the mainstream of the evolution of mathematics have contributed to the lack of recognition of his originality. As for clarity, while giving much credit to Peano, Constance Reid says of Hilbert that in the *Grundlagen der Geometrie* he [**20**, p. 60] "attempted to present the modern point of view with even greater clarity than either Pasch or Peano." What could be clearer than: "The reader may understand by the sign **1** any category whatever of entities"? Let the reader compare for himself the clarity of Dedekind's presentation of the foundations of arithmetic with that of Peano. There can be no doubt that the famous five axioms for the natural numbers are rightly called Peano's Postulates.

References

1. E. Beltrami, Saggio di interpretazione della geometria noneuclidea, Giorn. Mat. Battaglini, 6 (1868) 284–312.
2. E. W. Beth, The Foundations of Mathematics, North-Holland, Amsterdam, 1959.
3. G. Fano, Sui postulati fondamentali della geometria in uno spazio lineare a un numero qualunque di dimensioni, Giorn. Mat. Battaglini, 30 (1892) 106–132.
4. Hans Freudenthal, Die Grundlagen der Geometrie um die Wende des 19. Jahrhunderts, Math. -Phys. Semesterber., 7 (1961) 2–25.
5. ———, The main trends in the foundations of geometry in the 19th century, Logic, Methodology and Philosophy of Science, Stanford University Press, Stanford, California, 1962.
6. Jean van Heijenoort, ed., From Frege to Gödel, A source book in mathematical logic, 1879–1931, Harvard University Press, Cambridge, Mass., 1967.
7. David Hilbert, Grundlagen der Geometrie, Leibzig, 1899.
8. ———, The Foundations of Geometry, Trans. by E. J. Townsend, Open Court, La Salle Ill., 1902.
9. ———, Über den Zahlbegriff, Jber. Deutsch. Math.-Verein., 8 (1900) 180–184.
10. ———, Gesammelte Abhandlungen, Vol. 3, Springer-Verlag, Berlin, 1935.
11. P. E. B. Jourdain, The development of the theories of mathematical logic and the principles of mathematics, Quart. J. Pure Appl. Math., 43 (1912) 219–314.
12. H. C. Kennedy, The mathematical philosophy of Giuseppe Peano, Philos. Sci., 30 (1963) 262–266.
13. Moritz Pasch, Vorlesungen über neuere Geometrie, Teubner, Leibzig, 1882.
14. G. Peano, Arithmetices Principia, Nova methodo exposita, Bocca, Turin, 1889.
15. ———, I principii di Geometria, logicamente esposti, Bocca, Turin, 1889.
16. ———, Sul concetto di numero, Rivista di Matematica, 1 (1891) 87–102, 256–267.
17. ———, Sul § 2 del Formulario t. II: Aritmetica, Rivista di Matematica, 6 (1896–99) 75–89.
18. ———, Opere Scelte, Vol. 2, Edizioni Cremonese, Rome, 1958.
19. ———, Opere Scelte, Vol. 3, Edizioni Cremonese, Rome, 1959.
20. Constance Reid, Hilbert, Springer-Verlag, New York, 1970.
21. Hao Wang, The axiomatization of arithmetic, J. Symb. Logic, 22 (1957) 145–158.
22. R. L. Wilder, Evolution of Mathematical Concepts, Wiley, New York, 1968.

EMMY NOETHER

CLARK H. KIMBERLING, University of Evansville

The past two years have seen a surge of interest in Emmy Noether and her mathematics. Along with Auguste Dick's biography of her, listed below, Constance Reid's biography, *Hilbert*, frequently mentions Emmy Noether. New mathematics books, such as *Introduction to the Calculus of Variations*, by Hans Sagan, and *Commu-*

Clark Kimberling received his 1970 Ph. D. at Illinois Tech under A. Sklar. His research interests are real analysis and the history of mathematics. *Editor*.

tative Rings, by Irving Kaplansky, are spreading anew her methods, and the adjective "noetherian" abounds in titles to papers in mathematics research journals. The State University of New York at Buffalo has just set up a George William Hill–Emmy Noether Fellowship. A high school textbook, *Modern Introductory Analysis*, by Dolciani, Donnelly, Jurgensen, and Wooten, devotes a page to Emmy Noether. And one finds such remarks in periodical literature as "The woman mathematician today is better off than Emmy Noether, who taught without pay. But...".

Despite all this recent interest, it is difficult to find much about Emmy Noether in mathematics history books. Although she was dubbed "der Noether" by P. S. Alexandroff—and that name with its masculine German article has stuck, she is given only a footnote in E. T. Bell's *Men of Mathematics* and hardly more in comparable books. In fact, little else can be found about her than three obituary addresses and the biography published just last year:

(1) "Emmy Noether," by Hermann Weyl (memorial address at Bryn Mawr College, April 26, 1935), Scripta mathematica III, 3 (1935), pp. 201–220.
(2) "Nachruf auf Emmy Noether," by B. L. van der Waerden (in German), Mathematische Annalen 111 (1935), pp. 469–476.
(3) "Emmy Noether," by P. S. Alexandroff, address to the Moscow Mathematical Society, Sept. 5, 1935.
(4) "Emmy Noether," by Auguste Dick (in German), Birkhäuser Verlag, Basel (Switzerland), 1970.

Since (3) was not published I shall draw from it more than from the others.

1. Her passing. A note in the files of the *Bryn Mawr Alumnae Bulletin* reads, "The above was inspired, if not written, by Dr. Hermann Weyl, eminent German mathematician. Mr. Einstein had never met Miss Noether." The "above" is the following, as it appeared in *The New York Times*, May 3, 1935:

The efforts of most human beings are consumed in the struggle for their daily bread, but most of those who are, either through fortune or some special gift, relieved of this struggle are largely absorbed in further improving their worldly lot. Beneath the effort directed toward the accumulation of worldly goods lies all too frequently the illusion that this is the most substantial and desirable end to be achieved; but there is, fortunately, a minority composed of those who recognize early in their lives that the most beautiful and satisfying experiences open to humankind are not derived from the outside, but are bound up with the development of the individual's own feeling, thinking and acting. The genuine artists, investigators and thinkers have always been persons of this kind. However, inconspicuously the life of these individuals runs its course, none the less the fruits of their endeavors are the most valuable contributions which one generation can make to its successors.

Within the past few days a distinguished mathematician, Professor Emmy Noether, formerly connected with the University of Göttingen and for the past two years at Bryn Mawr College, died in her fifty-third year. In the judgment of the most competent living mathematicians, Fräulein Noether was the most significant creative mathematical genius thus far produced since the higher education of women began. In the realm of algebra,

in which the most gifted mathematicians have been busy for centuries, she discovered methods which have proved of enormous importance in the development of the present-day younger generation of mathematicians. Pure mathematics is, in its way, the poetry of logical ideas. One seeks the most general ideas of operation which will bring together in simple, logical and unified form the largest possible circle of formal relationships. In this effort toward logical beauty spiritual formulae are discovered necessary for the deeper penetration into the laws of nature.

Born in a Jewish family distinguished for the love of learning, Emmy Noether, who, in spite of the efforts of the great Göttingen mathematician, Hilbert, never reached the academic standing due her in her own country, none the less surrounded herself with a group of students and investigators at Göttingen, who have already become distinguished as teachers and investigators. Her unselfish, significant work over a period of many years was rewarded by new rulers of Germany with a dismissal, which cost her the means of maintaining her simple life and the opportunity to carry on her mathematical studies. Farsighted friends of science in this country were fortunately able to make such arrangements at Bryn Mawr College and at Princeton that she found in America up to the day of her death not only colleagues who esteemed her friendship but grateful pupils whose enthusiasm made her last years the happiest and perhaps the most fruitful of her entire career.

ALBERT EINSTEIN.

Princeton University, May 1, 1935.

2. Early years. We are indebted to Dr. Auguste Dick of Vienna for much of what we know today about Emmy Noether's early life and her forebears. Most of the information in this present section may be found in Dr. Dick's biography.

Among those affected by an 1809 Tolerance Edict in the German state of Baden was one Elias Samuel, who as the head of a Jewish household was required to change his name and the names of his nine children. He chose the surname *Nöther*, and one of his sons, Hertz, he renamed Hermann. At the age of eighteen, Hermann left his birthplace, Bruchsal, and went to Mannheim to study theology. However, in 1837, he and his older brother Joseph founded an iron-wholesaling firm. The firm lasted for nearly a century, when it fell to anti-Jewish forces.

Born to Hermann and Amalia Nöther were five children, and the third, in 1844, was named Max. During his fourteenth year, Max suffered from infantile paralysis and was somewhat handicapped for the rest of his life. Nevertheless, he became a mathematician of great stature, arriving at the University of Erlangen as a professor in 1875, where he remained until his death in 1921. In 1880, Max married Ida Amalia Kaufmann. Although their marriage certificate bears the name *Nöther*, Max and all his children used the name *Noether* instead.

Amalie Emmy Noether was born on March 23, 1882 in the South German town of Erlangen. She was the first child of Max and Ida Noether and soon had brothers, Albert, born in 1883, and Fritz, in 1884. Still another brother was born in 1889. The family rented a large flat in the first story of an apartment house at Nürnberger Strasse 30–32. Another tenant there for many years was Professor Eilhard Wiedemann, remembered as an Islamist as well as a physicist. The Noether family occupied their flat for about forty-five years.

As a child, Emmy was acutely near-sighted, not outwardly attractive, and not exceptional in any way. Her teachers and classmates remember that she favored the study of language and that little she did reflected teachings of the Jewish religion. Like many other girls, she took clavier lessons and dancing lessons, but apparently with little fervor.

Three years after leaving her "high school," the Städtischen Höheren Töchterschule in Erlangen, Emmy took tests for prospective schoolteachers of French and English. These tests were given in Ansbach in April, 1900. No sooner had she passed these and thus qualified as a language teacher than she became interested in university studies.

Among nearly a thousand students at the University of Erlangen in the winter of 1900, Emmy Noether was one of two women. As a rule, female students could not be registered in the usual sense, and they could take an examination for course credit only upon consent of the professor teaching the course. This consent was often withheld. Nevertheless, whether passing through the prerequisite courses in the usual manner or not, a woman could eventually take an examination for a university certificate.

Among Emmy's early professors at Erlangen, one was a historian and another, Julius Pirson, a professor of romance languages. Between 1900 and 1902, Emmy must have chosen to pursue mathematics rather than languages, since during that time she must have been preparing for the final university examination, which she passed in July, 1903. This examination was given in Nürnberg at the royal Realgymnasium, now the Willstätter-Gymnasium. Quite possibly it was administered by the mathematician Aurel Voß, from whom Emmy's brother Fritz later received his doctorate.

In the winter of 1903 Emmy attended classes at the University of Göttingen. There she heard such eminent mathematicians as Hermann Minkowski, Otto Blumenthal, Felix Klein, and David Hilbert. After just one semester, however, she returned to Erlangen, for it had become possible for women to be matriculated and tested in the manner formerly reserved for men.

In October of 1904, Emmy Noether was officially registered as a student at the University of Erlangen. As a member of Section II of the Philosophical Faculty, she studied only mathematics. On December 13, 1907, she passed her doctoral oral examination, and in July of 1908 her dissertation was registered with the Erlanger Universitätsschriften as Number 202.

3. Excerpts from Weyl's address. Concerning the dissertation and the professor, Paul Gordan, under whom Emmy wrote it, Weyl spoke as follows in his memorial address:

Side by side with [Max] Noether acted in Erlangen as a mathematician the closely befriended Gordan, an offspring of Clebsch's school like Noether himself. Gordan had come

to Erlangen shortly before, in 1874, and he, too, remained associated with that university until his death in 1912. Emmy wrote her doctor's thesis under him in 1907: "On complete systems of invariants for ternary biquadratic forms"; it is entirely in line with the Gordan spirit and his problems. The *Mathematische Annalen* contains a detailed obituary of Gordan and an analysis of his work, written by Max Noether with Emmy's collaboration. Besides her father, Gordan must have been well-nigh one of the most familiar figures in Emmy's early life, first as a friend of the house, later as a mathematician also; she kept a profound reverence for him though her own mathematical taste soon developed in quite a different direction. I remember that his picture decorated the wall of her study in Göttingen. These two men, the father and Gordan, determined the atmosphere in which she grew up. Therefore I shall venture to describe them with a few strokes.

Riemann had developed the theory of algebraic functions of one variable and their integrals, the so-called Abelian integrals, by a function-theoretic transcendental method resting on the minimum principle of potential theory which he named after Dirichlet, and had uncovered the purely topological foundations of the manifold function-theoretic relations governing this domain. (Stringent proof of Dirichlet's principle which seemed so evident from the physicist's standpoint was only given about fifty years later by Hilbert.) There remained the task of replacing and securing his transcendental existential proofs by the explicit algebraic construction starting with the equation of the algebraic curve. Weierstrass solved this problem (in his lectures published in detail only later) in his own half function-theoretic, half algebraic way, but Clebsch had introduced Riemann's ideas into the geometric theory of algebraic curves and Noether became, after Clebsch had passed away young, his executor in this matter: he succeeded in erecting the whole structure of the algebraic geometry of curves on the basis of the so-called Noether residual theorem. This line of research was taken up later on, mainly in Italy; the vein Noether struck is still a profusely gushing spring of investigations; among us, men like Lefschetz and Zariski bear witness thereto. Later on there arose, beside Riemann's transcendental and Noether's algebraic-geometric method, an arithmetical theory of algebraic functions due to Dedekind and Weber on the one side, to Hensel and Landsberg on the other. Emmy Noether stood closer to this trend of thought. A brief report on the arithmetical theory of algebraic functions that parallels the corresponding notions in the competing theories was published by her in 1920 in the *Jahresberichte der Deutschen Mathematikervereinigung*. She thus supplemented the well-known report by Brill and her father on the algebraic-geometric theory that had appeared in 1894 in one of the first volumes of the *Jahresberichte*. Noether's residual theorem was later fitted by Emmy into her general theory of ideals in arbitrary rings. This scientific kinship of father and daughter — who became in a certain sense his successor in algebra, but stands beside him independent in her fundamental attitude and in her problems — is something extremely beautiful and gratifying. The father was — such is the impression I gather from his papers and even more from the many obituary biographies he wrote for the *Mathematische Annalen* — a very intelligent, warm-hearted harmonious man of many-sided interests and sterling education.

Gordan was of a different stamp. A queer fellow, impulsive and one-sided. A great walker and talker — he liked that kind of walk to which frequent stops at a beer-garden or a cafe belong. Either with friends, and then accompanying his discussions with violent gesticulations, completely irrespective of his surroundings; or alone, and then murmuring to himself and pondering over mathematical problems; or if in an idler mood, carrying out long numerical calculations by heart. There always remained something of the eternal "Bursche" of the 1848 type about him — an air of dressing gown, beer and tobacco, relieved however by a keen sense of humor and a strong dash of wit. When he had to listen to others, in classrooms or at meetings, he was always half asleep. As a mathematician not of Noether's rank, and of an essentially different kind, Noether himself concludes his characterization of him with

the short sentence: "Er war ein Algorithmiker." His strength rested on the invention and calculative execution of formal processes. There exist papers of his where twenty pages of formulas are not interrupted by a single text word; it is told that in all his papers he himself wrote the formulas only, the text being added by his friends. Noether says of him: "The formula always and everywhere was the indispensable support for the formation of his thoughts, his conclusions and his mode of expression . . . . In his lectures he carefully avoided any fundamental definition of conceptual kind, even that of the limit."

He, too, had belonged to Clebsch's most intimate collaborators, had written with Clebsch their book on Abelian integrals; he later shifted over to the theory of invariants following his formal talent; here he added considerably to the development of the so-called symbolic method, and he finally succeeded in proving by means of this computative method of explicit construction the finiteness of a rational integral basis for binary invariants. Years later Hilbert demonstrated the theorem much more generally for an arbitrary number of variables — by an entirely new approach, the characteristic Hilbertian species of methods, putting aside the whole apparatus of symbolic treatment and attacking the thing itself as directly as possible. *Ex ungue leonem* — the young lion Hilbert showed his claws. It was, however, at first only an existential proof providing for no actual, finite algebraic construction. Hence Gordan's characteristic exclamation: "This is not mathematics, but theology!" What then would he have said about his former pupil Emmy Noether's later "theology", that abhorred all calculation and operated in a much thinner air of abstraction than Hilbert ever dared!

It is queer enough that a formalist like Gordan was the mathematician from whom her mathematical orbit set out; a greater contrast is hardly imaginable than between her first paper, the dissertation, and her works of maturity; for the former is an extreme example of formal computations and the latter constitute an extreme and grandiose example of conceptual axiomatic thinking in mathematics. Her thesis ends with a table of the complete system of covariant forms for a given ternary quartic consisting of not less than 331 forms in symbolic representation. It is an awe-inspiring piece of work; but today I am afraid we should be inclined to rank it among those achievements with regard to which Gordan himself once said when asked about the use of the theory of invariants: "Oh, it is very useful indeed; one can write many theses about it."

In 1910 Gordan retired, soon to be replaced by Ernst Fischer. In Weyl's judgment Fischer had a more penetrating influence on Emmy Noether's work than Gordan did. Weyl wrote as follows:

Under his direction the transition from Gordan's formal standpoint to the Hilbert method of approach was accomplished. She refers in her papers at this time again and again to conversations with Fischer. This epoch extends until about 1919. The main interest is concentrated on finite rational and integral bases; the proof of finiteness is given by her for the invariants of a finite group (without using Hilbert's general basis theorem for ideals), for invariants with restriction to integral coefficients, and finally she attacks the same question along with the question of a minimum basis consisting of independent elements for fields of rational functions.

4. Her contribution to physics. In 1916, Emmy Noether left Erlangen and went to the University of Göttingen. At that time Hilbert was working on the general theory of relativity and Emmy was especially welcome because of her knowledge of the theory of invariants.

Weyl described her major contribution to two important aspects of relativity as "the genuine and universal mathematical formulation: first, the reduction of the problem of differential invariants to a purely algebraic one by use of 'normal coordinates'; second, the identities between the left sides of Euler's equations of a problem of variation which occur when the (multiple) integral is invariant with respect to a group of transformations involving arbitrary functions (identities that contain the conservation theorem of energy and momentum in the case of invariance with respect to arbitrary transformations of the four world coordinates)."

During my own inquiries about Emmy Noether, it was once hinted that "young physicists are using her theories," and I was eventually referred to Professor Eugene Wigner (1963 Noble Prize in Physics), who wrote, "We physicists pay lip service to the great accomplishments of Emmy Noether, but we do not really use her work. Her contribution to physics that is most often quoted arose from a suggestion of Felix Klein. It concerns the conservation laws of physics, which she derived in a way which was at that time novel and should have excited physicists more than it did. However, most physicists know little else about her, even though many of us who have a marginal interest in mathematics have read much else by and about her."

Professor Peter G. Bergmann of Syracuse University gave the following account of Emmy Noether's influence in physics:

Noether's Theorem, so-called, forms one of the corner stones of work in general relativity as well as in certain aspects of elementary particles physics. The idea is, briefly, that to every invariance or symmetry property of the laws of nature (or of a proposed theory) there corresponds a conservation law, and vice versa. Accordingly, if a physical quantity is known to satisfy a conservation law (known as a "good quantum number" in quantum physics), the theorist attempts to construct a theory with appropriate symmetry properties. Conversely, if a theory is known to possess certain symmetries, then this fact alone entails the existence of certain integrals of the dynamical equations.

General relativity is characterized by the principle of general covariance, according to which the laws of nature are invariant with respect to arbitrary curvilinear coordinate transformations that satisfy minimal conditions of continuity and differentiability. A discussion of the consequences in terms of Noether's theorem (whether explicitly quoted as such or not) would have to include all of the work on ponderomotive laws, inter alia.

Goldstein's text, *Classical Mechanics*, contains a treatment of Noether's theorem on pps. 47 ff., without, however, calling it by that (or any other) name. J. L. Anderson's book, *Principles of Relativity Physics* (Academic Press, 1967) explicitly refers to Noether's Theorem on p. 92. These references, picked at random from my book shelves at home, will indicate to you that a list of papers involving Noether's theorem in one way or other would probably amount to hundreds of items.

5. World War I years. At Göttingen it was still difficult, as it had been in Erlangen, for anyone to push through any provision for remuneration for Dr. Noether. The philologists and historians of the Göttingen Philosophical Faculty opposed Hilbert's efforts in her behalf, and Hilbert once declared during a University Senate meeting, "I do not see that the sex of the candidate is an argument against her ad-

mission as *Privatdozent*. After all, we are a university and not a bathing establishment." Finally, in 1919, her habilitation as *Privatdozent* was made possible, and three years later she became a "nicht-beamteter ausserordentlicher Professor," under which title she received no salary. A small salary was soon afforded her, however, as a lecturer in algebra.

Weyl's description of Emmy Noether's political life is interesting as a commentary on pre-World War II Germany:

During the wild times after the Revolution of 1918, she did not keep aloof from the political excitement, she sided more or less with the Social Democrats; without being actually in party life she participated intensely in the discussion of the political and social problems of the day. One of her first pupils, Grete Hermann, belonged to Nelson's philosophic-political circle in Göttingen. It is hardly imaginable nowadays how willing the young generation in Germany was at that time for a fresh start, to try to build up Germany, Europe, society in general, on the foundations of reason, humaneness, and justice. But alas! the mood among the academic youth soon enough veered around; in the struggles that shook Germany during the following years and which took on the form of civil war here and there, we find them mostly on the side of the reactionary and nationalistic forces. Responsible for this above all was the breaking by the Allies of the promise of Wilson's Fourteen Points, and the fact that Republican Germany came to feel the victors' fist not less hard than the Imperial Reich could have; in particular, the youth were embittered by the national defamation added to the enforcement of a grim peace treaty. It was then that the great opportunity for the pacification of Europe was lost, and the seed sown for the disastrous development we are the witnesses of. In later years Emmy Noether took no part in matters political. She always remained, however, a convinced pacifist, a stand which she held very important and serious.

6. Excerpts from Alexandroff's address. Emmy Noether's mathematical activities from 1919 to 1923 and her influence on the mathematical community are covered by Alexandroff in his 1935 address to the Moscow Mathematical Society:

Emmy Noether entered upon her wholely individual path of mathematical work in 1919–1920. She herself dated the beginning of this principal period of activity with the well-known collaborative work with V. Schmeidler (*Mathematische Zeitschrift*, vol. 8, 1920). This work serves as a prologue to her general theory of ideals, opening with the classical memoir of 1921, *Idealtheorie in Ringbereiche*. I think that of all that Emmy Noether did, the bases of the general theory of ideals and all the work related to them have exerted, and will continue to exert, the greatest influence on mathematics as a whole.... If the development of today's mathematics undoubtedly proceeds under the aegis of algebra, and algebraic concepts and algebraic methods have penetrated into the various mathematical theories themselves, then all that has become possible only after the works of Emmy Noether. She taught us just to think in simple, and thus general, terms: homomorphic representation, the group or ring with operators, the ideal—and not in complicated algebraic calculations, and she therefore opened a path to the discovery of algebraic regularities where before these regularities had been obscured by complicated specific conditions.

It is enough to glance at the work of Pontryagin in the theory of continuous groups, at the just completed work of Kolmogoroff in the combinatorial topology of locally-bicompact spaces, at the works of Hopf in the theory of continuous representations, not to mention the works of van der Waerden in algebraic geometry, to feel the influence of Emmy Noether's

ideas. This influence is vividly clear also in the book by H. Weyl, *Gruppentheorie und Quantenmechanik*.

For all the concreteness and constructiveness of Emmy Noether's various findings, as related to the various working periods of her life, there is no doubt that her greatest energy and the major thrust of her talent were directed toward general mathematical conceptions which had to be axiomatically tinctured to a considerable degree. It is quite appropriate to analyze this aspect of her work in more detail—especially because now the question of general and specific, abstract and concrete, axiomatic and constructive, appears as one of the most acute questions of mathematical practice. Interest in the problem as a whole is sharpened by the fact that, on one hand, mathematical journals are, without doubt unnecessarily, burdened with an abundance of all sorts of generalizing, axiomatic, and similar articles, often devoid of concrete mathematical content; while on the other hand, here and there declarations are heard that only that which is "classical" comprises the true mathematics. Under this latter slogan, important mathematical problems are rejected only because they oppose one or another habit of thought, or because they employ concepts that were not current several decades ago.... H. Weyl, in the obituary that I have already cited, also raises this general question. What he says in this regard penetrates so far into the heart of the matter that I cannot but quote him in full.

"In a conference on topology and abstract algebra as two ways of mathematical understanding, in 1931, I said this:

"Nevertheless I should not pass over in silence the fact that today the feeling among mathematicians is beginning to spread that the fertility of these abstracting methods is approaching exhaustion. The case is this: that all these nice general notions do not fall into our laps by themselves. But definite concrete problems were first conquered in their undivided complexity, singlehanded by brute force, so to speak. Only afterwards the axiomaticians came along and stated: Instead of breaking in the door with all your might and bruising your hands, you should have constructed such and such a key of skill, and by it you would have been able to open the door quite smoothly. But they can construct the key only because they are able, after the breaking in was successful, to study the lock from within and without. Before you can generalize, formalize and axiomatize, there must be a mathematical substance. I think that the mathematical substance in the formalizing of which we have trained ourselves during the last decades, becomes gradually exhausted. And so I foresee that the generation now rising will have a hard time in mathematics."

"Emmy Noether," H. Weyl continues, "protested against that: and indeed, she could point to the fact that the axiomatic method in her hands had opened new, concrete, profound problems and pointed the way to their solution."

In this quotation there is much that deserves attention: First of all, of course, the indisputable point of view that a concrete, I would say, naïve, seizure of mathematical material must precede any axiomatic treatment of it; that, further, the axiomatic treatment is only of interest when it touches upon real mathematical knowledge (the "mathematical substance," of which H. Weyl speaks), and does not appear, to speak crudely, as a milling of the wind. All this is indisputable, and it is not against this that Emmy Noether protested. But she did protest against that pessimism which is seen in the last words cited by Weyl himself from his speech of 1931; the substance of human knowledge, including mathematical knowledge, is inexhaustible, at least for many long years to come—in this Emmy Noether firmly believed. The "substance of the *last decades*" is exhausting itself, but not mathematical substance in general, which by a thousand complicated threads is connected with the reality of the world's and mankind's existence. Emmy Noether intensely felt this connection of every great mathematical system, even the most abstract, with real existence, and even if she did not think this connection out philosophically, she felt it with the whole being of a learned, lively

person, who was by no means shackled within abstract schemes. For Emmy Noether mathematics was always knowledge of the world and not a game of symbols, and she avidly protested when representatives of those areas of mathematics which are immediately concerned with applications wanted to secure privilege for practical knowledge.

In 1924–25 the school of Emmy Noether made one of its most brilliant acquisitions: a graduating Amsterdam student, B. L. van der Waerden, became her pupil. He was then 22 years old and one of the brightest young mathematical talents in Europe. Van der Waerden quickly mastered the theories of Emmy Noether, enlarged them with important new findings, and like no one else, promoted her ideas. A course in the general theory of ideals, given by van der Waerden in 1927 in Göttingen, was enormously successful. The ideas of Emmy Noether in the brilliant exposition of van der Waerden subdued public mathematical opinion, first at Göttingen, then in the other leading mathematical centers of Europe. It was no accident that Emmy Noether required a popularizer of her ideas: her lectures were intended for a small group of students, working in the direction of her own investigations and listening constantly to her. From external appearances, Emmy Noether's delivery was poor, hurried, and inconsistent; but in her lectures there was immense strength of mathematical thought and extraordinary animation and fervor. Of such a kind, too, were her reports to mathematical societies and at meetings. For the mathematician who had already been captured by her ideas and become interested in her work, her reports provided much; but the mathematician who stood far from her work often could understand her exposition only with great difficulty.

From 1927 the influence of the ideas of Emmy Noether on contemporary mathematics continually grew, and along with it grew scientific praise for the author of those ideas. The direction of her work at this time moved more and more into the region of non-commutative algebra, the theory of representation and of the general arithmetic of hypercomplex areas. Two fundamental works of the last period of her activity are *Hyperkomplexe Grossen und Darstellungstheorie* (1929) and *Nichtcommutative Algebra* (1933), both published in *Mathematische Zeitschrift* (vols. 30 and 37). These and related works evoked considerable response from spokesmen for the algebraic theory of numbers, especially from Helmut Hasse. Among her pupils during this period of her activities, the most outstanding was M. Deuring; in addition there was a whole row of young, beginning mathematicians (Witt, Fitting, and others).

Emmy Noether at last received recognition for her ideas. If in the years 1923–25 she had to demonstrate the importance of the theories that she had developed, in 1932, at the International Mathematical Congress in Zürich, she was crowned with the laurel of her success. A summary of her work read by her at this gathering was the real triumph of the direction she represented, and she could look, not only with inner satisfaction, but now also with consciousness of full recognition, upon the mathematical path that she had traveled. The Zürich congress was the high point of her international scientific reputation. In a few months there would burst over German culture, and in particular over her home, which the University of Göttingen had become, the catastrophe of the Fascist revolution, which in a few weeks scattered to the wind all that had been built up over a long period of decades. One of the greatest tragedies that human culture has undergone since the time of the Renaissance took place, a tragedy which a few years ago appeared improbable and impossible in Europe of the 20th century. One of its numerous victims was the Göttingen School of Algebra, which had been founded by Emmy Noether. Its directress was banished from the walls of the University; and having lost the right to teach, Emmy Noether had to emigrate from Germany. She accepted the invitation from the women's college at Bryn Mawr, where she lived out the last year and a half of her life.

If what I have just quoted is the main strand of the material of Alexandroff's address, another is his description of Emmy Noether's influence on Soviet mathematics and her regard for Soviet ideals:

Emmy Noether was closely connected with Moscow. This connection began in 1923, when the late Pavel Samuelovitch Urysohn and I first arrived in Göttingen and immediately found ourselves in a mathematical circle whose leader was Emmy Noether. The basic traits of the Noether school struck us right away: the scientific enthusiasm of the directress of the school which was passed along to all her students, her deep belief in the importance and mathematical fruitfulness of her ideas (a belief that was not at all shared then by everyone, even in Göttingen), and the extraordinary simplicity and sincerity of relations between the head of the school and its members. In those days this school was almost entirely made up of young Göttingen students; the time was still in the future when it would become, for its membership and for its acknowledged world-wide influence, an outstanding international center of algebraic thought.

The mathematical interests of Emmy Noether (centered at that time in the full swing of her work on the general theory of ideals) and the mathematical interests of Urysohn and myself (centered around the problems of so-called abstract topology) had many points in common and quickly led to continual, almost daily, mathematical discussions. Emmy Noether was interested, however, not only in our topological work, but also in what had been taking place in the whole area of mathematics (and not only in the area of mathematics) in Soviet Russia; she did not hide her sympathies with our country and our social and governmental system, in spite of the fact that the manifestation of these sympathies seemed outrageous and unseemly to the majority of representatives of Western European academic circles. The matter had reached the point where Emmy Noether was literally banished from one of the Göttingen boarding houses (where she had settled and lived) at the demand of the student corporation, resident in the same house, who did not want to live under the same roof with a "Marxist-inclined Jewess."

And Emmy Noether was truly gladdened by the scientific, and particularly the mathematical successes of the Soviet country, since she saw in this the final refutation of all the old wives' tales to the effect that "the Bolsheviks are destroying culture." A spokesman of the most abstract areas of mathematical science, she distinguished herself at the same time by a surprising sensitivity in understanding the great historical movements of our epoch; always vitally interested in politics, hating war with her whole being, and hating chauvinism in all its manifestations, she never in this area knew any vacillation: her sympathies always and unchangingly belonged to the Soviet Union, in which she saw the beginning of a new era in the history of mankind and firm support for everything progressive for which human thought has lived and lives still.

The scientific and personal friendship which sprang up between Emmy Noether and me in 1923 did not come to an end even with her death. Recalling this friendship in his obituary speech, Weyl advances the supposition that the general system of thought of Emmy Noether did not remain without influence on my own topological research. I am happy now to affirm the truth of Weyl's supposition: Emmy Noether's influence on my own, and on other topological research in Moscow, was very great, and it affected the whole essence of our work. In particular, my theory of the continuous breakdown of topological spaces arose to a significant degree under the influence of conversations with her in December and January of 1925–26, when we were in Holland together.

Emmy Noether spent the winter of 1928–29 in Moscow. She taught a course in abstract algebra at the University of Moscow and conducted a seminar in algebraic geometry at

the Communist Academy. She quickly established contact with a majority of Moscow's mathematicians, in particular and especially, with L. S. Pontryagin and O. U. Schmidt. It is not difficult to trace the influence of Emmy Noether on the mathematical talent of L. S. Pontryagin; a strong algebraic note in his work was undoubtedly benefited in its development by contact with Emmy Noether. In Moscow, Emmy Noether very easily fit herself in with our life, both in her scientific and her non-professional relationships. She lived in a modest room in the KSU hostel near the Crimean Bridge, and most of the time she walked to the University. She was very much interested in the life of our country, especially in the life of Soviet young people and students.

In the winter of 1928–29 I was as usual on a visit to Smolensk and was giving lectures on algebra at the Pedagogical Institute there. Inspired by my continual conversations with Emmy Noether, I gave my lectures along the lines established by her. Among my students there, A. G. Kurosh immediately stood out, and the theories that I was expounding, wholely steeped as they were in Emmy Noether's ideas, appealed very much to him. In this way, through my teaching, Emmy Noether acquired a disciple who has since grown into an independent and learned man, as is well known, and whose works through the present day have proceeded in the principal circle of ideas created by her.

In the spring of 1929, she left Moscow for Göttingen with the firm intention of coming to visit us again in the near future. Several times she was close to carrying out that intention, and closest to doing so in the last year of her life. After her exile from Germany, she seriously considered a final trip to Moscow, and I exchanged letters with her in this regard. She clearly understood that nowhere could she find the means to create a new brilliant mathematical school in exchange for the one that had been taken from her in Göttingen. I had already conducted talks with the Narkompros [The People's Commissariat for Education] about assigning her a chair at the University of Moscow. However, at the Commissariat, as usual, they were slow in making a decision, and they did not give me a final answer. Meanwhile, time passed, and Emmy Noether, deprived even of that modest work which she had in Göttingen, could wait no longer and had to accept the invitation of the women's college...

Such was Emmy Noether, the greatest of women mathematicians, a great scientist, an amazing teacher, and an unforgettable person....True, Weyl has said that "the Graces did not stand at her cradle," and he is right, if one has in mind the generally known heaviness of her appearance. But here Weyl is speaking of her not only as a great scholar, but also as a great woman. And she was that—her femininity appeared in that gentle and subtle lyricism which lay at the heart of the far-flung but never superficial concerns which she maintained for people, for her profession, and for the interests of all mankind. She loved people, science, life, with all the warmth, all the cheerfulness, all the unselfishness, and all the tenderness of which a deeply sensitive—and feminine—soul is capable.

7. In America. Among the scientists who left Germany during the early thirties and sooner or later took refuge in the United States were E. Artin, R. Courant, P. Debye, M. Dehn, A. Einstein, P. Ewald, W. Feller, J. Franck, K. Friedrichs, K. Gödel, E. Hellinger, O. Neugebauer, J. von Neumann, Emmy Noether, L. Nordheim, O. Ore, G. Pólya, G. Szegö, A. Tarski, Olga Taussky (Todd), H. Weyl, and E. Wigner.

Arrangements were made for Emmy Noether to teach at Bryn Mawr College, just outside Philadelphia, beginning in the autumn of 1933. Conveniently close was the Institute for Advanced Study at Princeton, where, beginning in February, 1934, she gave weekly lectures. At the Institute were Einstein, Weyl, Oswald Veblen, and Abraham Flexner.

Emmy Noether returned to Germany to visit during the summer of 1934 and then resumed her work at Bryn Mawr and Princeton in the early fall. Richard Brauer had joined the Institute, and after her lectures, she usually visited with Brauer, Weyl, and Veblen before returning to Bryn Mawr.

One of Emmy Noether's associates at Bryn Mawr was Grace Shover (Quinn), now a professor at the American University. Awarded the Emmy Noether Fellowship for post-doctoral study, she became acquainted with Emmy Noether in September, 1934. There were three other graduate students in mathematics. Marie Weiss of Newcomb College held the Emmy Noether Scholarship. Olga Taussky held the foreign fellowship. Ruth Stauffer (McKee) was a doctoral candidate, Emmy Noether's only American Ph.D. student.

Professor Quinn recalls that Emmy Noether "was around 5′4″ tall and was slightly rotund in build. Her complexion was swarthy. Her dark hair, flecked with grey, was cropped short. She wore thick glasses to cover her near-sighted eyes, and she had a way of turning her head aside and looking into the distance when trying to think while talking. Her looks and dress were most unconventional, such as to attract attention, but such a result was far from her thoughts. She was sincere, straightforward, kindly, thoughtful, and considerate.

"Her lectures were delivered in broken English. She often lapsed into her native German when she was bothered by some idea in lecturing.

"She loved to walk. She would take her students off for a jaunt on a Saturday afternoon. On these trips she would become so absorbed in her conversation on mathematics that she would forget about the traffic and her students would need to protect her."

The chairman of the Bryn Mawr mathematics department was Anna Pell Wheeler, now deceased. Having studied at Göttingen a few years before receiving her doctorate from the University of Chicago in 1910, Professor Wheeler became a very close friend of Emmy Noether. In this connection, Mrs. McKee has written, "Probably the greatest difference in her life in America was her close friendship with the head of the mathematics department. In Germany at that time women were neither expected nor encouraged to study. It was rather assumed that their role in life was that of a homemaker. Therefore, to have as a friend a woman who was a nationally recognized mathematician who had earlier studied at Göttingen and who thoroughly understood the problems of a woman scholar in Germany, was a unique experience for Miss Noether.... Many of Miss Noether's former students and colleagues stopped to see [her], in Bryn Mawr and she always took them to see her 'good friend.'"

8. Missing letters recovered. Together with Jean Cavaillés, Emmy Noether edited the correspondence between Richard Dedekind and Georg Cantor. Although completed in March, 1933, the book, *Briefwechsels—G. Cantor und R. Dedekind*, did not appear until 1937, when it was published by Hermann of Paris.

The Cantor-Dedekind letters were still in Emmy Noether's possession when she

died, along with correspondence from G. Fröbenius and H. Weber (Hilbert's predecessor at Göttingen). A representative of the law firm which settled Emmy Noether's estate wrote to her brother Fritz, exiled in Tomsk, asking what should be done with the letters. Professor Noether's reply was that they be returned to their (unspecified) owner. This directive was not carried out. Instead, the letters lay lost in the files of a Philadelphia law office, until, after some 33 years, a member of the firm wrote me, "Inasmuch as you are researching her life, a rather valuable bit of information was unearthed by me in going through the Estate file. Under separate cover you will shortly receive them via parcel post. I suppose you will be agreeable to the modest charge of $25.00..."

I had no idea what the "valuable bit of information" was, but as promised, the letter collection arrived a few days later. Included with the famous Cantor-Dedekind letters, a few of whose paragraphs may be found in English in Sherman K. Stein's popular *Mathematics: The Man-made Universe* (Freeman), are 47 letters written by Weber in Königsberg and Heidelberg to Dedekind in Braunschweig. Together with 20 post cards, telegrams, and printed circulars, these letters span the years 1876–79.

Three letters each by Fröbenius and Dedekind are dated 1882–83. Their remaining 38 letters and 7 postcards are dated from 1895 to 1901. Most were written in 1886. Their content is more mathematical than that of the Weber letters. A few reach a length of 20 pages. I counted a total of 178 pages from Fröbenius to Dedekind and 113 from Dedekind to Fröbenius.

At present, the letters are kept in the Clifford Memorial Library at the University of Evansville.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

ON THE FUNDAMENTAL PROBLEM OF MATHEMATICS

P. ERDÖS, Hungarian Academy of Science

I read with interest the paper of C. Goffman, "And What is your Erdös Number?" (this MONTHLY, 76 (1969) 791). For some time I have considered a problem which I feel is of much more fundamental importance.

We define a graph $\mathfrak{G}(M)$ as follows: the vertices of our graph $\mathfrak{G}(M)$ are the mathematicians. Two vertices are joined if the corresponding mathematicians have written at least one joint paper. (For the time being, let us ignore the papers with more than two authors.)

Is $\mathfrak{G}(M)$ planar, that is, can it be imbedded in E^2? I was not able to solve this interesting and important question. It seems that $\mathfrak{G}(M)$ does not, at present, contain a complete pentagon $\mathfrak{K}(5)$. It certainly contains a $\mathfrak{K}(4)$; for example, in the set Erdös-Rényi-Szekeres-Turan, each pair has a joint paper.

I communicated this problem to Schinzel, who proved that $\mathfrak{G}(M)$ is not planar by showing that $\mathfrak{G}(M)$ contains a $\mathfrak{K}(3,3)$ — that is, a complete bipartite graph of 6 vertices (with three vertices of each color and the 9 edges connecting black to white in all possible ways). The white vertices are Chowla, Mahler, Schinzel; the black ones are Davenport, Erdös, Lewis; the simple task of finding the 9 relevant papers can be left to the reader.

I would like to mention some interesting related problems. There are sets of three mathematicians, each subset of which has a paper (more precisely, only the empty set has no papers); for example, Erdös-Rogers-Taylor. It would be nice to have an example of a set of 4 mathematicians where each of the 15 non-empty subsets has a paper. I believe such a set does not yet exist.

The graph $\mathfrak{G}(M)$, in fact, should be denoted by $\mathfrak{G}^{(t)}(M)$ (t stands for time). I suggest the following optimistic conjecture: to each integer r there is a time t_r so that for $t > t_r$ the graph $\mathfrak{G}^{(t)}(M)$ contains a complete graph $\mathfrak{K}(r)$ of r vertices.

INITIAL DIGITS FOR THE SEQUENCE OF PRIMES

R. E. Whitney, Lock Haven State College

The initial digit problem is concerned with the frequency of occurrence of elements with initial digit $a \in \{1, \cdots, 9\}$ in a sequence of positive integers. In this discussion, the sequence of positive integers with initial digit a will be denoted by $A = \{a_\nu\}$. Also, P will denote the sequence of primes.

It is well known that the logarithmic density of A in the sequence of positive integers is $\log_{10}(1 + 1/a)$, see [**1**]. The purpose of this note is to show that the relative logarithmic density of A in P is also $\log_{10}(1 + 1/a)$. This is an unusual result because of the irregular distribution of the primes. As a consequence of this result, one might say that 1 is the preferred initial digit for the sequence of primes.

The relative logarithmic density, $d(A)$, of A in P will be defined as

$$d(A) = \lim_{x\to\infty} \left[\sum_{\substack{a_\nu \leq x \\ a_\nu \in P\cap A}} 1/a_\nu \Big/ \sum_{\substack{p \leq x \\ p\in P}} 1/p \right]. \tag{1}$$

As usual, the upper and lower relative logarithmic densities, $\overline{d(A)}$ and $\underline{d(A)}$, are obtained by replacing 'limit' in (1) by 'limit superior' and 'limit inferior' respectively.

Since

$$\sum_{\substack{p \leq x \\ p\in P}} 1/p = \log\log x + B_1 + O(1/\log^2 x), \tag{2}$$

where B_1 is a constant, [2] an equivalent expression for $d(A)$ is

$$d(A) = \lim_{x\to\infty}\Bigl[\sum_{\substack{a_\nu \leqq x\\ a_\nu\in P\cap A}} 1/a_\nu\Bigr]\Big/\log\log x. \tag{3}$$

Using (2), we have .

$$\sum_{\substack{a10^t<p\leqq(a+1)10^t\\ p\in P}} 1/p = \log\frac{\log(a+1)10^t}{\log a10^t} + O(1/t^2) \qquad (t \geqq 1). \tag{4}$$

Thus

$$\begin{aligned}\sum_{\substack{a_\nu\ (a+1)10^n\\ a_\nu\in P\cap A}} 1/a_\nu &= \sum_{t=0}^{n}\ \sum_{a10^t<p\leqq(a+1)10^t} 1/p\\ &= \log\Bigl[\prod_{t=0}^{n}\frac{B_2+t}{A_2+t}\Bigr] + O\Bigl(\sum_{t=0}^{n} 1/t^2\Bigr),\end{aligned} \tag{5}$$

where $B_2 = \log_{10}(a+1)$ and $A_2 = \log_{10} a$. Evidently

$$\begin{aligned}\underline{d(A)} &\geqq \lim_{n\to\infty}\Bigl[\sum_{\substack{a_\nu\leqq a10^n\\ a_\nu\in P\cap A}} 1/a_\nu\Bigr]\Big/\log\log a10^n\\ &\geqq \lim_{n\to\infty}\Bigl[\sum_{\substack{a_\nu\leqq(a+1)10^{n-1}\\ a_\nu\in P\cap A}} 1/a_\nu\Bigr]\Big/\log\log a10^n\\ &= \lim_{n\to\infty}\Bigl[\log\prod_{t=0}^{n-1}\frac{B_2+t}{A_2+t}\Bigr] + O(1)/\log\log 10 + \log(n+A_2).\end{aligned} \tag{6}$$

If we apply the usual Euler formula [3] for the Gamma Function, viz.

$$\Gamma(z) = \lim_{n\to\infty}\frac{n!\,n^z}{\prod_{i=0}^{n}(z+i)}, \tag{7}$$ then (6) becomes

$$\begin{aligned}\underline{d(A)} &\geqq \lim_{n\to\infty}\frac{\log[\Gamma(A_2)/\Gamma(B_2)] + (B_2-A_2)\log(n-1) + O(1)}{\log(n+A_2)+\log\log 10}\\ &= B_2 - A_2 = \log_{10}(1+1/a).\end{aligned} \tag{8}$$

Similarly,

$$\begin{aligned}\overline{d(A)} &\leqq \lim_{n\to\infty}\Bigl[\sum_{\substack{\nu\nu\leqq(a+1)10^n\\ a_\nu\in P\cap A}} 1/a_\nu\Bigr]\Big/\log\log a10^n\\ &= \lim_{n\to\infty}\frac{\log[\Gamma(A_2)/\Gamma(B_2)] + (B_2-A_2)\log n + O(1)}{\log(n+A_2)+\log\log 10}.\end{aligned} \tag{9}$$

Thus

(10) $\overline{d(A)} \leqq \log_{10}(1 + 1/a)$ and the desired conclusion,

(11) $d(A) = \log_{10}(1 + 1/a)$, follows.

The above result can be generalized to cover any specified sequence of initial digits. If the sequence of positive integers with initial digit sequence $\{a_1, a_2, \cdots, a_n\}$ is denoted by $A(a_1, a_2, \cdots, a_n)$, then the relative logarithmic density of $A(a_1, a_2, \cdots, a_n)$ in P is

$$\log_{10}\left(1 + 1 \Big/ \sum_{i=1}^{n} 10^{n-i} a_i\right).$$

Thus of all the specified sequences of initial digits of length n, in the primes, the preferred initial sequence is 10^{n-1}.

References

1. R. L. Duncan, A note on the initial digit problem, Fibonacci Quart. No. 5, 7 (1969) 474–5.
2. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen I, 1909, 200–1. Chelsea, New York, 1953.
3. E. T. Whittaker and G. N. Watson, A Course of Modern Analysis, 4th, ed., Cambridge. Univ. Press, New York, 1961, 237.

ANOTHER PROOF OF A RESULT OF PERRY ON CHAINS OF FINITE SETS

D. J. KLEITMAN, Massachusetts Institute of Technology

and MORDECHAI LEWIN, Technion–Israel Institute of Technology, Haifa.

1. Introduction. In [1] Harzheim proves the following theorem:

THEOREM 1. *Let n be a nonnegative integer. Then there exists a positive integer $N(n)$ such that for any set A of $N(n)$ elements and any mapping f of the set of nonempty subsets of A into A such that $f(X) \in X$ for all $X \subseteq A$, there exists a strictly increasing sequence*

$$X_0 \subset X_1 \subset \cdots \subset X_n \subseteq A$$

such that $f(X_0) = f(X_1) = \cdots = f(X_n)$.

In [2] Perry shows that $N(n)$ may always be taken as 2^n. It is the purpose of this note to supply a different proof of Perry's result. The resulting method does not apply for Rado's generalization of Harzheim's result [3], in which a subset is mapped onto a non-empty subset of fixed size (or not exceeding a prescribed size) contained in it and not necessarily onto an element.

2. The proof. Let S be a set. $|S|$ denotes the cardinal of S. The obliteration

operator ^ serves to remove from any system of elements the element above which it is placed.

Let $S = \{1,2,\cdots,N\}$ and let f be an arbitrary, but fixed, choice function defined on the set of nonempty subsets of S. For $1 \leqq i \leqq N$ let $\mathfrak{S}_i$ be the size of a maximal chain of subsets of S whose image is i. Perry's result may now be stated in the following form:

THEOREM 2. *If $N \geqq 2^k$, then for some i, $1 \leqq i \leqq N$, $\mathfrak{S}_i \geqq k$.*

This theorem is a direct consequence of the following lemma:

LEMMA. *If $N = 2^k + j$, $j \leqq 2^k$, then*

$$\sum_{i=1}^{N} (\mathfrak{S}_i - k + 1) \geqq 2j + 1. \tag{1}$$

To see this we only have to note that the R.H.S. of (1) is positive. Then clearly one of the terms of the left side of (1) is positive, which is the theorem.

For a subset S' of S, $\mathfrak{S}_i'$ $(i \in S')$ will refer to subsets of S' alone.

(1) may be put in the following form:

$$\sum_{i=1}^{N} \mathfrak{S}_i - N(k-1) \geqq 2j + 1. \tag{1'}$$

The lemma is true for $k = 1, j = 0$. We therefore assume it true for $j - 1, j > 0$ and k, and prove it for j, k. Consider $S' = \{1,2,\cdots,N-1\}$. Then by our hypothesis

$$\sum_{i=1}^{N-1} \mathfrak{S}_i' - (N-1)(k-1) \geqq 2j - 1 > 0. \tag{2}$$

Then for some i_0, $1 \leqq i_0 \leqq N - 1$ we have

$$\mathfrak{S}_{i_0}' > k - 1. \tag{3}$$

Consider now $S'' = \{1,2,\cdots,N;\hat{\imath}_0\}$. Again by hypothesis

$$\sum_{i \neq i_0} \mathfrak{S}_i'' - (N-1)(k-1) \geqq 2j - 1. \tag{4}$$

Adding i_0 to S'' and noting that $\mathfrak{S}_i' \leqq \mathfrak{S}_i$, $\mathfrak{S}_i'' \leqq \mathfrak{S}_i$ for every i and that the set S itself increases the length of the maximal chain with image $f(S)$, we may write

$$\sum_{i=1}^{N} \mathfrak{S}_i - N(k-1) = 1 + \sum_{i \neq i_0} \mathfrak{S}_i'' + \mathfrak{S}_{i_0}' - (N-1)(k-1) - (k-1). \tag{5}$$

Rearranging the right side of (5) and using (3) and (4) we obtain (1).

For $j = 2^k$, we have $N = 2j$ and (1) becomes

$$\sum_{i=1}^{N} \mathfrak{S}_i - Nk \geqq 1, \tag{6}$$

which is the lemma for $k+1$ and $j=0$. This proves the lemma and the theorem.

Supported in part by NSF grant GP-22928.

References

1. E. Harzheim, Ein kombinatorisches Problem über Auswahlfunktionen, Publ. Math. (Debrecen), 15 (1968) 19–22.
2. R. L. Perry, Representatives of subsets, J. Comb. Theory, 3 (1967) 302–304.
3. R. Rado, A theorem on chains of finite sets, J. London Math. Soc., 42 (1967) 101–106.

SOME DECOMPOSITIONS OF THE INTEGERS FROM 0 TO $p^n - 1$

S. W. GOLOMB, University of Southern California

1. Introductory and historical notes. From the two-digit decimal notation, it is clear that every integer from 0 to 99 has a unique representation as a sum $a+b$, where a is in the set $\{0,1,2,3,4,5,6,7,8,9\}$ and b is in the set $\{0,10,20,30,40,50,60,70,80,90\}$. We shall be concerned with the most general decompositions of this type. We shall see that for prime bases, only a simple generalization of the n-digit representation in base p can occur, but that for composite bases there are stranger arrangements which also arise.

While these results are not covered in many standard books, two papers of de Bruijn [1] and [2] lead back to a fundamental article by Redei [3], which uses cyclotomic methods and implicitly encompasses the results of this paper. However, Redei's presentation is far less accessible to the non-specialist. Further work along these lines was published by Sands [4], and [5] covers somewhat similar ground.

The theorems in this paper were first formulated as empirical conjectures during the investigation of the problem of "non-standard counters", as described in [6].

2. The decomposition theorems.

THEOREM 1. *Suppose there are two sets of numbers, $A=\{a_1,a_2,\cdots,a_p\}$ and $B=\{b_1,b_2,\cdots,b_p\}$ such that the p^2 numbers $\{a_i+b_i\}$ are all distinct modulo p^2. If p is prime, then one of the two sets (A or B) consists of a complete residue system modulo p, while the other set consists, modulo p^2, of the numbers $\{c+ip\}$, $i=0,1,\cdots,p-1$, for some integer c, $0 \leqq c \leqq p-1$.*

Proof. Let $\zeta = e^{2\pi i/p^2}$, a primitive p^2-root of unity. Let

$$\alpha = \sum_{a_i \in A} \zeta^{a_i} \quad \text{and} \quad \beta = \sum_{b_j \in B} \zeta^{b_j}.$$

Then

$$\alpha\beta = \sum_{a_i \in A} \zeta^{a_i} \sum_{b_j \in B} \zeta^{b_j} = \sum_{A,B} \zeta^{a_i+b_j} = \sum_{k=1}^{p^2} \zeta^k = 0,$$

and since α and β are two complex numbers whose product is 0, at least one of them

(without loss of generality, suppose α) is 0. Then ζ is a root of the polynomial equation

$$a(x) = \sum_{a_i \in A} x^{a_i} = 0$$

of degree $\leqq p^2 - 1$. However, the minimal polynomial for ζ is the cyclotomic polynomial

$$\Phi_{p^2}(x) = 1 + x^p + x^{2p} + \cdots + x^{(p-1)p},$$

which is irreducible of degree $p^2 - p$. Then $a(x)$ must be of the form $g(x) \cdot \Phi_{p^2}(x)$, where the extra factor $g(x)$ has degree $\leqq p-1$. Since $\Phi_{p^2}(x)$ already has p terms, with exponents spaced p apart, multiplication by $g(x)$ will increase the number of terms, violating the definition of $a(x)$, unless $g(x)$ is a monomial, say $g(x) = x^c$ with $0 \leqq c \leqq p-1$. In this case,

$$a(x) = g(x)\Phi_{p^2}(x) = x^c + x^{c+p} + x^{c+2p} + \cdots + x^{c+(p-1)p},$$

and the set A consists of $\{c, c+p, c+2p, \cdots, c+(p-1)p\}$, as required, modulo p^2.

Given that A is of this form, we consider the set $\{a_i + b_j\}$ which reduces modulo p to $\{c + b_j\}$, so that if $\{a_i + b_j\}$ is to take on all distinct values modulo p^2, it is clearly *necessary* for b_j to take on all distinct values modulo p.

We observe that it is also *sufficient* to take $\{a_i\} = \{c + ip\}$, $i = 0, 1, \cdots, p-1$, and $\{b_j\} = \{$any complete residue system modulo $p\}$, in order for $\{a_i + b_j\}$ to assume the p^2 distinct values modulo p^2.

Examples.

1. For any integer n, the sets $A = \{0, n, 2n, \cdots, (n-1)n\}$ and $B = \{0, 1, 2, \cdots, n-1\}$ produce as sums $a_i + b_j$, with $a_i \in A$ and $b_j \in B$, all the numbers from 0 to $n^2 - 1$, in what is essentially their base n representation. Theorem 1 indicates that for *prime* n, only the obvious modifications of this construction are possible.

2. For the composite case, the following example for $n = 4$ is typical: $\{0, 1, 8, 9\} + \{0, 2, 4, 6\} = \{0, 1, 2, 3, \cdots, 15\}$, in a way which is basically different from the construction of Theorem 1. For this case, since $\Phi_{16}(x) = 1 + x^8$, *any* polynomial $x^c(1 + x^d)(1 + x^8)$ with $1 \leqq d \leqq 7$ and $0 \leqq c \leqq 7-d$ will generate a set of four exponents which may be used as the set A.

3. More generally, since $\Phi_{2r}(x)$ divides $1 + x^r$, for even $n = 2m$ we have $\Phi_{n^2}(x)$ divides $1 + x^{2m^2}$. If $h(x)$ is a sum of m distinct powers of x, all lower than the $2m^2$ power, then $h(x)[1 + x^{2m^2}]$ has n distinct exponents which may be used as the set A.

A similar argument holds for *all* composite values of n.

Next we prove the n-dimensional generalization of Theorem 1, which was merely the two-dimensional case.

THEOREM 2. *Suppose that $A_1, A_2, \cdots, A_n$ are n sets of p integers each, p prime,*

such that the p^n sums $\sum_{i=1}^{n}\alpha_i$ with $\alpha_i \in A_i$ are all distinct modulo p^n. Then the n sets, with appropriate reordering, can be described as follows: The elements of A_i, taken modulo p^i, are the numbers $\{c_i + jp^{i-1}\}$, for a fixed integer c_i, and $j = 0, 1, 2, \cdots, p-1$.

Proof. We use induction on n. For $n = 1$, $A = A_1$ must itself be a complete residue system modulo p, so that we may take $c_1 = 0$ and use $A = \{j\}, 0 \leqq j \leqq p-1$.

For $n = 2$, the assertion reduces to Theorem 1, which has already been proved separately.

Assume the assertion is known to hold for all $n < N$, and consider the case of N sets. Let $\eta = e^{2\pi i/p^N}$, and form

$$\prod_{i=1}^{N}\left(\sum_{\alpha_i \in A_i}\eta^{\alpha_i}\right) = \sum_{j=1}^{p^N}\eta^j = 0,$$

so that at least one of the factors in the product (e.g., the N-th factor) must equal 0. Then

$$\sum_{\alpha \in A_N}\eta^{\alpha} = 0,$$

and the polynomial

$$u(x) = \sum_{\alpha \in A_N} x^{\alpha}$$

must be a multiple of

$$\Phi_{p^N}(x) = 1 + x^{p^{N-1}} + x^{2p^{N-1}} + \cdots + x^{(p-1)p^{N-1}}.$$

Since $u(x)$ has degree $\leqq p^N - 1$ and has only p terms, we must have $u(x) = x^c\Phi_{p^N}(x)$, whereby $A_N = \{c + jp^{N-1}\}$ with $0 \leqq j \leqq p-1$ as required. Modulo p^{N-1}, the elements of A_N are all congruent, which necessitates that $A_1 + A_2 + \cdots + A_{N-1}$ must generate all p^{N-1} distinct values modulo p^{N-1}. By the inductive hypothesis, the sets $A_1, A_2, \cdots, A_{N-1}$ must then have the required form.

General Notes.

1. Theorem 2 describes a form of n-digit representation in base p for the numbers from 0 to $p^n - 1$.

2. The fact that $\Phi_{p^n}(x)$ is the sum of p distinct powers of x, and that these cyclotomic polynomials are irreducible over the rational numbers, plays a central role in the proofs of the theorems just given. There does not seem to be any proof which is more "elementary."

3. As with Theorem 1, many more things can happen if p is composite. Again, the same algebraic methods can be used to explore these possibilities.

This research was supported in part by the Air Force Office of Scientific Research under Grant No. AFOS–68–1555–C.

References

1. N. G. de Bruijn, On the factorization of finite abelian groups, Indag. Math., Nederl. Akad. Wetensch., 15 (1953), 258–264.
2. N. G. de Bruijn, On the factorization of cyclic groups, Indag. Math., Nederl. Akad. Wetensch., 15 (1953) 370–377.
3. L. Redei, Ein Beitrag zum Problem der Faktorisation von endlichen abelschen Gruppen, Acta. Math. Hung., Vol. 1, (1950) 197–206.
4. A. D. Sands, On the factorisation of finite Abelian groups, Acta Math. Hungar., 8 (1957) 65–68.
5. L. Carlitz and L. Moser, On some special factorizations of $(1-x^n)/(1-x)$, Canad. Math. Bull., No. 4, 9 (1966) 421–426.
6. M. Cohn, S. Even, S. W. Golomb, and A. Lempel, The stability of counting sequences under stage delays, SIAM J. Appl. Math., No. 2, 20 (March 1971) 183–188.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

IDENTITIES ON MATRICES

KIRBY C. SMITH AND HILLEL J. KUMIN, University of Oklahoma

Let S_k be the group of all permutations on $\{1,2,3,\cdots,k\}$. Let $A_1, A_2, \cdots, A_k$ be any $n \times n$ matrices with entries from a field. We define

$$[A_1,\cdots,A_k] = \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(k)},$$

where $\operatorname{sgn}(\sigma) = \pm 1$ depending on whether σ is an even or odd permutation. For example if $k = 2$ we have

$$[A_1, A_2] = A_1A_2 - A_2A_1,$$

and if $k = 3$

$$[A_1, A_2, A_3] = A_1A_2A_3 - A_1A_3A_2 + A_3A_1A_2 - A_3A_2A_1 + A_2A_3A_1 - A_2A_1A_3.$$

The following theorem was proved by Amitsur and Levitzki [1] using only elementary techniques but the proof was rather involved.

THEOREM. *If $A_1, A_2, \cdots, A_{2n}$ are any $n \times n$ matrices with entries from a field, then $[A_1, A_2, \cdots, A_{2n}] = 0$.*

It is easy to see that $2n$ is the least possible, for if E_{ij} is the $n \times n$ matrix with 1 in the i,jth position and zeroes elsewhere, then

$$[E_{11}, E_{12}, E_{22}, E_{23}, \cdots, E_{nn}] = E_{1n} \neq 0.$$

So for any $k < 2n$ we can find $n \times n$ matrices $A_1, A_2, \cdots, A_k$ such that $[A_1, A_2, \cdots, A_k] \neq 0$.

Amitsur and Levitzki's theorem can be stated in graph theory terms. Swan [3] has given a fairly simple and quite elementary proof of the theorem using graph theory techniques.

It is natural to restrict the class of all $n \times n$ matrices to some subclass and see if an analogue of the theorem can be obtained. For example, if we restrict ourselves to the class of all $n \times n$ symmetric matrices it is easy to see that $2n$ is still the best possible. For the class of skew-symmetric matrices, however, the answer does not seem to be the same. This leads us to our two conjectures.

CONJECTURE 1. *Let $A_1, A_2, \cdots, A_{2n-2}$ be any $n \times n$ skew-symmetric matrices, then $[A_1, A_2, \cdots, A_{2n-2}] = 0$.*

CONJECTURE 2. *If $k < 2n - 2$ then there are $n \times n$ skew-symmetric matrices $A_1, A_2, \cdots, A_k$ such that $[A_1, A_2, \cdots, A_k] \neq 0$.*

Kostant [2] gave a third proof of the theorem of Amitsur and Levitzki using advanced techniques. He was also able to answer Conjecture 1 in the affirmative when n is even using cohomology theory. The case where n is odd is still unsettled as far as we know. Nothing seems to have been done on Conjecture 2.

The theorem of Amitsur and Levitzki was an early result in the study of algebras satisfying a polynomial identity (PI-algebras) and an answer to Conjectures 1 and 2 may lead to results which parallel the uses of that theorem.

References

1. S. Amitsur and J. Levitzki, Minimal identities for algebras, Proc. Amer. Math. Soc., 1 (1950) 449–463.

2. B. Kostant, A theorem of Frobenius, a theorem of Amitsur-Levitzki, and cohomology theory, J. Math. Mech., 7 (1958) 237–264.

3. Richard G. Swan, An application of graph theory to algebra, Proc. Amer. Math. Soc., 14 (1963) 367–373.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

ON INVOLUTIONS OF A CIRCLE

W. F. PFEFFER, University of California, Davis

Many easily expressible geometric theorems about spheres often require proofs which involve rather sophisticated techniques of algebraic topology and thus are completely inaccessible to the undergraduate students. The purpose of this note is to give an elementary proof of the existence of a coincidence point of two free involutions of a circle. Equivalently stated, we shall prove that if two involutions of a circle have no fixed points, their composition always has a fixed point. This is a fixed point theorem which does not follow from Lefschetz degree considerations.

In a complex plane C we shall consider the unit circle $S^1 = \{z \in C : |z| = 1\}$. A continuous map $\sigma: S^1 \to S^1$ is called an **involution** of S^1 if $\sigma^2 = \sigma \circ \sigma$ is the identity map of S^1. An involution σ of S^1 is called **free** if it has no fixed points, i.e., if $\sigma(x) \neq x$ for all $x \in S^1$. The **antipodal** map $\alpha: S^1 \to S^1$ defined by $\alpha(z) = -z$ is an example of a free involution. If $z_1, z_2 \in C$, we shall denote by (z_1, z_2) the open line segment connecting z_1 and z_2, i.e., $(z_1, z_2) = \{tz_1 + (1-t)z_2 : 0 < t < 1\}$.

LEMMA. *Let σ be a free involution of S^1. Then*

$$(x, \sigma(x)) \cap (y, \sigma(y)) \neq \varnothing$$

for every $x, y \in S^1$.

Proof. Choose $x \in S^1$. Since $\sigma(x) \neq x$, $S^1 - \{x, \sigma(x)\} = A \cup B$ where A and B are open connected arcs. Because σ is a continuous map and σ^2 is the identity, either $\sigma(A) = A$ or $\sigma(A) = B$ must hold. If $\sigma(A) = A$ then also

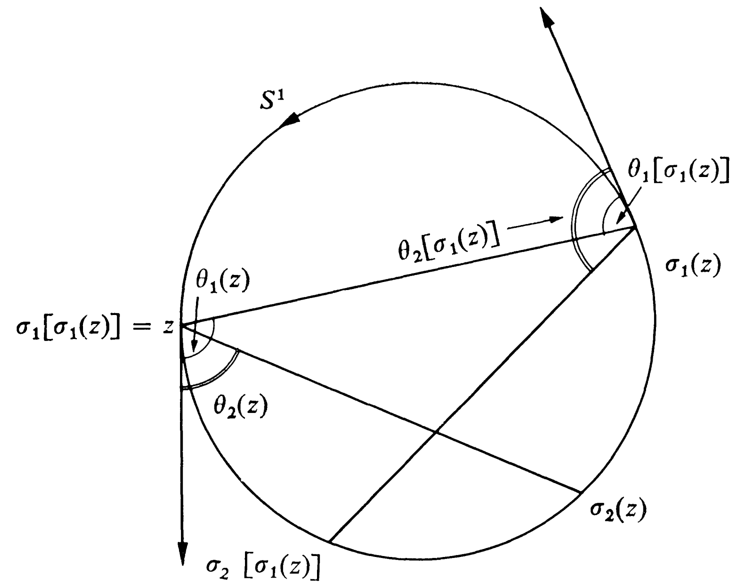
$$\sigma(A \cup \{x, \sigma(x)\}) = A \cup \{x, \sigma(x)\}.$$

But $A \cup \{x, \sigma(x)\}$ is homeomorphic to the closed unit interval $[0, 1]$ and thus there is a $y \in A \cup \{x, \sigma(x)\}$ such that $\sigma(y) = y$; but this is a contradiction. Therefore $\sigma(A) = B$ and the lemma follows.

PROPOSITION. *Let σ_1 and σ_2 be two free involutions of S^1. Then there is a point $z \in S^1$ such that $\sigma_1(z) = \sigma_2(z)$.*

Proof. The angle between two vectors determined by non-zero complex numbers x and y is defined in the usual way as a length of the smaller arc of S^1 determined by points $x/|x|$ and $y/|y|$. If $z \in S^1$ and $i = 1, 2$, we denote by $\theta_i(z)$ the angle between the vectors determined by $\sigma_i(z) - z$ and $z \cdot \sqrt{-1}$. Clearly the function $\theta: S^1 \to (-\pi, \pi)$

defined by $\theta(z) = \theta_1(z) - \theta_2(z)$, $z \in S^1$, is continuous. Choose $z \in S^1$. If $\theta(z) = 0$ then, of course, $\sigma_1(z) = \sigma_2(z)$. Thus without loss of generality we may assume that $\theta(z) > 0$ (see the picture). Since $(z, \sigma_2(z)) \cap (\sigma_1(z), \sigma_2[\sigma_1(z)]) \neq \emptyset$, we have $\theta[\sigma_1(z)] < 0$. From the intermediate value theorem it follows that $\theta(z_0) = 0$ for some $z_0 \in S^1$, and the proof is completed.



COROLLARY. *Let σ be an involution of S^1. Then there is a point $z \in S^1$ such that either $\sigma(z) = z$ or $\sigma(z) = -z$.*

This corollary follows immediately from the previous proposition applied to σ and the antipodal map α.

Note. The above proposition is also valid for higher dimensional spheres. This is deduced, e.g., in [1], 33.6, page 89, as a simple corollary from the generalized Borsuk-Ulam theorem. However, the proof of the generalized Borsuk-Ulam theorem is itself quite intricate and makes extensive use of algebraic topology.

Reference

1. P. E. Conner and E. E. Floyd, Differentiable Periodic Maps, Springer Verlag, Berlin, 1964.

MAXIMA AND MINIMA OF FUNCTIONS OF TWO VARIABLES

MICHEL NICOLA, Shimer College

It is generally believed that the techniques used to investigate local extrema of functions of one variable are not adequate for settling similar problems in several variables. Thus if every section of the surface representing $f(x, y)$ by a vertical plane

through one point on the surface is a curve with a minimum value at that point, this does not guarantee that $f(x, y)$ itself has a minimum value at that point; the usual example is $f(x, y) = (y - x^2)(y - 2x^2)$. When restricted to any line through the origin of the xy-plane it has a minimum at zero, yet the function does not have a minimum at the origin; it is negative when $x^2 < y < 2x^2$. Nevertheless, by imposing additional conditions it is possible to develop this method into one applicable to functions of several variables. Indeed, all cases which can be handled by the usual tests are covered by our method as well as some cases when the usual tests are indecisive.

In the following we limit ourselves to functions of two variables, although the results can be generalized without difficulty to functions of any number of variables. All the maxima and minima discussed are understood to be *local.* We imagine that we have performed a translation of coordinates so that the point at which we wish to test for an extremum is at the origin.

Let the transformation $T:(t, u) \to (x, y)$, with domain the rectangle $R(1) = \{(t, u): |t| < 1, |u| \leqq 1\}$, be defined by

(1) $$x = tu, \quad y = t(1 - u^2)^{\frac{1}{2}}.$$

The range of T is the open unit disk centered at the origin of the xy-plane, but T is not 1-1, as can be seen from the formula (1) defining y. For any t_1 such that $0 < t_1 \leqq 1$, let $R(t_1)$ be the rectangle

(2) $$R(t_1) = \{(t, u): |t| < t_1, |u| \leqq 1\}.$$

Then the restriction of T to $R(t_1)$ has as its range the open disk

(3) $$D(t_1) = \{(x, y): x^2 + y^2 < t_1^2\}.$$

If f is a real-valued function whose domain contains an open disk $D(t_1)$ defined by (3), let the function $g: R(t_1) \to \boldsymbol{R}$, where $\boldsymbol{R}$ denotes the real number field, be defined by

(4) $$g(t, u) \equiv f[tu, t(1 - u^2)^{\frac{1}{2}}], \quad |t| < t_1, \quad |u| \leqq 1.$$

For each u such that $|u| \leqq 1$, let the function $h_u: t \to h_u(t)$ be defined for $|t| < t_1$ by

(5) $$h_u(t) \equiv g(t, u).$$

If, for a given u, the function h_u has an mth derivative at t, then the value of this mth derivative at t will be denoted by $h_u^{(m)}(t)$. Finally, if this mth derivative exists for $|t| < t_1$ and $|u| \leqq 1$, let the function $p_m: R(t_1) \to \boldsymbol{R}$ be given by

(6) $$p_m(t, u) \equiv h_u^{(m)}(t).$$

Our method proceeds by investigating the extremum properties of $h_u(t)$ at $t = 0$

for each u with $|u| \leqq 1$. If, for example, $h_u^{(2)}(0) \neq 0$ for at least one u, then $f(0,0)$ is a minimum value of f provided $h_u^{(2)}(0) > 0$ for all u and p_2 is continuous throughout $R(t_1)$. Thus for $f(x,y) = \frac{1}{2}(x^2 + xy + y^2)$, a quick computation gives $h_u^{(2)}(0) = 1 + u(1-u^2)^{\frac{1}{2}}$, which is positive for all u; hence f has a minimum value at the origin.

If $h_u^{(2)}(0)$ is positive for some values of u and negative for other values, then $f(0,0)$ is neither a maximum nor a minimum value of f. If, however, $h_u^{(2)}(0)$ vanishes for at least one value of u and is positive for all other values of u, the test is indecisive unless $p_2(t,u)$ has a minimum value at $(0,u)$ for each u for which $h_u^{(2)}(0) = 0$, in which case $f(0,0)$ is a minimum value of f. And it might be possible to use the same test to test p_2 for minimum values, as in

EXAMPLE 1. $f(x,y) = x^2 - 3x^2y + y^4$.

The first partials as well as $f_{xy}^2 - f_{xx}f_{yy}$ vanish at the origin, so that the usual test fails. Our method gives

$$h_u(t) = u^2t^2 - 3u^2(1-u^2)^{\frac{1}{2}}t^3 + (1-u^2)^2t^4,$$

$$h_u^{(2)}(0) = 2u^2,$$

so that $h_u^{(2)} = 0$ for $u = 0$ and $h_u^{(2)} > 0$ for all $u \neq 0$. We should therefore test

$$p_2(t,u) \equiv h_u^{(2)}(t) = 2u^2 - 16u^2(1-u^2)^{\frac{1}{2}}t + 12(1-u^2)^2t^2$$

for a minimum value at the origin. For this purpose, let $u = ws$ and $t = (1-w^2)^{\frac{1}{2}}s$. Then

$$q_w(s) \equiv p_2(t,u) = 2(6-5w^2)s^2 + j_w(s),$$

where $j_w(s)$ is of order higher than the second in s.

$$q_w^{(2)}(0) = 4(6-5w^2) > 0$$

for all w with $|w| \leqq 1$. Therefore $p_2(0,0)$ is a minimum value of p_2; hence $f(0,0)$ is a minimum value of f.

If $k > 2$ is the lowest integer for which $h_u^{(k)}(0) \neq 0$ for at least one value of u, then all the above still applies provided we use $h_u^{(k)}(0)$ and p_k in place of $h_u^{(2)}(0)$ and p_2.

EXAMPLE 2. $f(x,y) = x^4 + x^3y + x^2y^2 + xy^3$.

Here the first and second partial derivatives vanish at $(0,0)$, and it is not quite clear how by the usual methods one can use the higher partial derivatives. Our method gives

$$h_u(t) = [u^2 + u(1-u^2)^{\frac{1}{2}}]t^4.$$

It is clear that $k = 4$ and

$$h_u^{(4)}(0) = 24[u^2 + u(1-u^2)^{\frac{1}{2}}],$$

which is positive for positive u and negative for $u = -0.6$. Hence $f(0,0)$ is neither a minimum nor a maximum value of f.

THEOREM 1. *Consider a function f with the properties specified above. Let the functions g, h_u, and p_m be as in* (4), (5), *and* (6). *Let k, where $k \leqq m$, be the smallest positive integer for which $h_u^{(k)}(0) \neq 0$ for at least one value of u. Then $f(0,0)$ is*

(i) *a local minimum (maximum) value of f if k is even and $h_u^{(k)}(0) > 0$ ($h_u^{(k)}(0) < 0$) for all u, and if p_k is continuous throughout $R(t_1)$,*

(ii) *neither a minimum nor a maximum value of f if k is odd, or if k is even and there exist at least two numbers a and b in the closed interval $[-1,1]$ such that $h_a^{(k)}(0) > 0$ while $h_b^{(k)}(0) < 0$.*

Proof. Let (x,y) be an arbitrary point of $D(t_1)$, and let (t,u) be a point of $R(t_1)$ such that (1) holds. Then, by Taylor's formula, there exists a θ such that $0 < \theta < 1$ and

$$h_u(t) = h_u(0) + (1/k!)t^k h_u^{(k)}(\theta t). \tag{7}$$

Therefore, from (5), (4), and (7) we get

$$h_u(t) - h_u(0) = f(x,y) - f(0,0) \equiv \Delta f = (1/k!)t^k h_u^{(k)}(\theta t). \tag{8}$$

We first consider (i); it clearly suffices to consider the minimum test only. Since p_k is continuous throughout $R(t_1)$ and $p_k(0,u) > 0$ for all u, there exists a t_2 satisfying $0 < t_2 \leqq t_1$ such that $p_k(t,u) > 0$ throughout $R(t_2)$. Therefore for all u and t with $|t| < t_2$ we have $h_u^{(k)}(t) > 0$; hence $h_u^{(k)}(\theta t) > 0$. It follows from (8) that $\Delta f \geqq 0$ throughout $D(t_2)$, and so $f(0,0)$ is a minimum value of f.

To prove (ii), suppose k is odd and, for definiteness, $h_u^{(k)}(0) > 0$ for $u = c$. Then since $h_c^{(k)}(t)$ exists for all t satisfying $|t| < t_1$, it is a continuous function of t. Hence there exists a t_3, where $0 < t_3 \leqq t_1$, such that for all t satisfying $|t| < t_3$ we have $h_c^{(k)}(t) > 0$, and therefore $h_c^{(k)}(\theta t) > 0$. It follows from (8) that Δf has the same sign as t. Thus for any t' such that $0 < t' < t_3$, we have $\Delta f > 0$ for those points on $D(t')$ for which $t > 0$ and $\Delta f < 0$ for the points on $D(t')$ for which $t < 0$, from which the result follows.

If k is even and $h_a^{(k)}(0) > 0$ while $h_b^{(k)}(0) < 0$, then an argument similar to the above shows that on any $D(t')$ we have $\Delta f > 0$ for those points for which $u = a$ and $\Delta f < 0$ for the points for which $u = b$. Hence $f(0,0)$ is neither a minimum nor a maximum value of f.

THEOREM 2. *Under the hypotheses of Theorem* 1, *let k be even and A be a proper subset of the closed interval $[-1,1]$ such that $h_u^{(k)}(0)$ vanishes for each $u \in A$ and is positive (negative) for each $u \notin A$. Then $f(0,0)$ is*

(i) *a local minimum (maximum) value of f if p_k is continuous throughout $R(t_1)$ and has a local minimum (maximum) value at $(0,u)$ for each $u \in A$,*

(ii) *neither a minimum nor a maximum value of f if for at least one $u \in A$ we find that $h_u(0)$ is not a minimum (maximum) value of h_u.*

Proof. We again present only the proof in the case of testing for a minimum value of f.

(i) The properties of p_k imply that there exists a t_4, where $0 < t_4 \leqq t_1$, such that $p_k(t, u) \geqq 0$ throughout $R(t_4)$. Therefore for all u and t with $|t| < t_4$ we have $h_u^{(k)}(t) \geqq 0$, and hence $h_u^{(k)}(\theta t) \geqq 0$. It follows from (8) that $\Delta f \geqq 0$ throughout $D(t_4)$, and the result follows.

(ii) Since $h_u(0)$ is not a minimum value of h_u for $u = a \in A$, it must be a maximum value of h_u, or else neither a maximum nor a minimum value. Thus there exists a t_5, where $0 < t_5 \leqq t_1$, such that $h_a^{(k)}(t) < 0$ for all t satisfying $0 < t < t_5$, or for all t satisfying $-t_5 < t < 0$, or for all t satisfying either of these two conditions; the same holds for $h_a^{(k)}(\theta t)$. But if $u = b \notin A$, there exists a t_6, where $0 < t_6 \leqq t_1$, such that $h_b^{(k)}(t) > 0$ for all t satisfying $|t| < t_6$; the same holds for $h_b^{(k)}(\theta t)$. We then see from (8) that on any $D(t')$, where $t' \leqq \inf\{t_5, t_6\}$, we have $\Delta f < 0$ for some points for which $u = a$ and $\Delta f > 0$ for all points for which $u = b$ and $t \neq 0$. The result follows from the different signs of Δf.

NOTE. In the remaining cases, when $h_u(0)$ is a minimum (maximum) value of h_u for each u and still $p_k(0, u)$ is not a minimum (maximum) value of p_k for at least one $u \in A$, the test fails.

I am indebted to Professor Eldon Boes and the fellow participants in the 1970 NSF Summer Institute at the New Mexico State University Department of Mathematics for many illuminating discussions and counterexamples.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

ACCREDITATION AND CERTIFICATION

Report to the Board of Governors of the Mathematical Association of America from its *Ad Hoc Committee to Consider Accreditation and Certification in Mathematics*

Background. In August 1968, the Board of Governors of the Mathematical Association of America asked CUPM to study the question of accreditation and certification in mathematics and to report its findings to the Board. The task of conducting that study and preparing a report was assigned to the CUPM Panel on College Teacher Preparation. The Panel's "Report on Accreditation and Certification" was accepted by CUPM and conveyed to the Board in the fall of 1969 for consideration at its January, 1970 meeting.

PROBLEMS AND SOLUTIONS

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before May 31, 1972. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2337. *Proposed by A. W. Walker, Toronto, Canada*

Show how to locate eleven coplanar points on eleven straight lines, with each point on three lines and three points on each line, using (a) straightedge and compasses; (b) straightedge only.

E 2338. *Proposed by A. W. Walker, Toronto, Canada*

Straight lines AP, BP, CP meet the side lines BC, CA, AB of triangle ABC at points D, E, F. By Euclidean construction, locate P so that it lies on the radical axis of circles ABC and DEF.

E 2339. *Proposed by A. W. Walker, Toronto, Canada*

Points D, E, F are the feet of the perpendiculars to the sides of triangle ABC from a point $P(\neq A, B, \text{ or } C)$ in the plane of the triangle. Prove that P cannot lie on the radical axis of circles ABC and DEF. (Cf. Problem E 2338.)

E 2340. *Proposed by Franz Hering, University of Washington*

A square matrix is **doubly stochastic** if its entries are nonnegative and if every row sum and every column sum is one. Show that every doubly stochastic matrix (other than the one with all entries equal) contains a 2×2 submatrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

such that either $\min(a,d) > \max(b,c)$ or $\max(a,d) < \min(b,c)$.

E 2341. *Proposed by Harry Lass, Jet Propulsion Laboratory, California Institute of Technology*

Given n urns numbered $1, 2, \cdots, n$ and k objects. Suppose that each of the objects is placed at random in one of the urns. For $r = 1, 2, \cdots, n$ let E_r be the event that the number of objects in the first r urns does not exceed r. Find the probability of the joint occurrence of $E_1, E_2, \cdots, E_n$ (Cf. E 2252 [1971, 797].)

E 2342. *Proposed (independently) by Joe Buhler, Reed College, and by M. B. Nathanson, University of Rochester*

If k and n are positive integers, what is the highest power of 2 that divides $k^n - 1$? In particular, for a fixed k, find all values of n for which $k^n \equiv 1 \pmod{2^n}$.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Functional Equation

E 2280 [1971, 196]. *Proposed by Felix Magnotta, Washington and Jefferson College*

Solve the functional equation

$$f(x+y) = f(x-y) + y[f'(x+y) + f'(x-y)].$$

Solution by Leon Gerber, St. John's University. Suppose that f satisfies the equation and let $g(x) = f(x) - f(0) - xf'(0)$. Then g also satisfies the equation and $g(0) = g'(0) = 0$. For $x = y = z/2$, we have $2g(z) = zg'(z)$, the solution of which is seen to be $g(z) = az^2$ where a is any constant. It follows then that every solution must be a quadratic: $f(x) = ax^2 + bx + c$. But obviously every quadratic satisfies the equation.

Several solvers differentiated the equation twice to show that $f'''(x) = 0$, so that f must be a quadratic. Many forgot to justify this step.

Also solved by seventy-one other readers.

A Special Leech Construction

E 2281 [1971, 196]. *Proposed by Cornelius Groenewoud, Snyder, New York*

Let Q be the midpoint of the line segment PR. Construct with compass and straightedge a triangle ABC having P for orthocenter, Q for incenter, and R for centroid.

Solution by Robin Robinson, Dartmouth College. It is assumed that $PR \neq 0$, i.e., that the triangle is not equilateral. The line PR is the Euler line, and the incenter lies on the Euler line only when the triangle is isosceles, in which case PR is an altitude. If S is the circumcenter, the four points P, Q, R, S are equally spaced on the Euler line at intervals of, say, $2d$. A routine application of analytic methods to the isosceles triangle with vertices at $(a,0)$, $(-a,0)$, $(0,c)$ shows that, if Q is to be equidistant from the three sides, then $c = a\sqrt{15}$ and $c = 15d$, with P: $(0,d)$, Q:$(0,3d)$, R: $(0,5d)$, S: $(0,7d)$, and $8d$ as radius of the circumscribed circle. The triangle is then unique, and is constructed as follows: Extend PQ beyond P by half its length, determining the point O. Erect the perpendicular to PQR at O; this is the base of the triangle, with PQR as altitude. With S as center, describe a circle of radius twice PR, cutting the base and the altitude at the required vertices.

Also solved by Leon Bankoff, Walter Bluger, Cal Poly Solution Group, R. G. Cassie, Jordi Dou (Spain), Leon Gerber, M. G. Greening (Australia), John Leech (Scotland), Simeon Reich (Israel), K. R. S. Sastry (Ethiopia), Wolfe Snow, Charles Wexler, Richard Yates, and the proposer.

Leech calls attention to his article on the general problem of constructing a triangle given its circumcenter, orthocenter, and incenter: see *An impossible construction*, Math. Gazette 38 (1954), 117–118.

Twelve-Tone Intervals

E 2283 [1971, 297]. *Proposed by Irving Adler, North Bennington, Vermont*

Composers using the twelve-tone scale have found that for any partition of the scale into two six-tone sets A and B, the musical intervals separating pairs of tones in B are the same as the musical intervals separating pairs of tones in A, and each interval has the same multiplicity in both sets. Consider the set of integers modulo $2n$ $(Z/2n)$. Partition this set into two sets A and B of n integers each. Show that the set of all differences including multiplicity (taken mod $2n$) is the same in each set.

Solution by William McWorter, Jr., Ohio State University. We prove the following: Suppose that Q is a finite quasigroup of order $2n$. [Thus every element of Q appears once and only once in each row and in each column in the multiplication table for Q; that is, the multiplication table for Q forms a Latin square. —Ed.] Partition Q such that $Q = A \cup B = C \cup D$, where $|A| = |B| = |C| = |D| = n$. Let $x \in Q$ be arbitrary. Then the number $N(x)$ of ways that x can be written in the form $x = ac$ with $a \in A$ and $c \in C$ is exactly the same as the number of ways that x can be written in the form $x = bd$ with $b \in B$ and $d \in D$.

To prove this, write the multiplication table of Q as below:

	C	D
A	$N(x)$	$n - N(x)$
B	$n - N(x)$	$N(x)$

The numbers in each block indicate the number of occurrences of x in that block. There are $N(x)$ occurrences in AC by definition. Since x occurs once and only once in each row, it follows that there are $n - N(x)$ occurrences in AD. Since x occurs once and only once in each column, there are $n - (n - N(x)) = N(x)$ occurrences in BD.

To solve the problem take $Q = Z/2n$ and let C be the set of (additive) inverses of elements in A, so that D is the set of inverses of elements in B.

Also solved by W. O. Alltop, P. H. Anderson, A. K. Austin (England), E. D. Bolker, D. W. Bouwsma, Cal Poly Solution Group, L. Carlitz & R. A. Scoville, Don Coppersmith, R. J. Dickson, J. R. Doner, R. C. Entringer, Bennington Gill, M. G. Greening (Australia), C. V. Heuer & G. A. Heuer, James Inglis, D. E. Knuth, A. G. Konheim, R. P. Kopp, H. C. Kranzer, Harry Lass, Douglas Lind, James Long, Carolyn MacDonald, E. P. McCravy, J. G. Mauldon, Joseph Pasciak, V. S. Poythress & H. S. Sun, Simeon Reich (Israel), G. B. Robinson, D. W. Roeder, J. Schonheim (Israel), L. E. Shader, David Spear, Stephen Spindler, D. J. Sterling, John Stout, D. P. Sumner, E. Szekeres (Australia), Konrad Victor (Israel), W. G. Wild, Gideon Yuval, Thomas Zaslavsky, and D. A. Zave.

Two Inequalities

E 2284 [1971, 297]. *Proposed by A. W. Walker, Toronto, Canada*

If a, b, c are positive numbers and if $x = (b+c-a)$, $y = (c+a-b)$, $z = (a+b-c)$, show that $abc\sum yz \geqq xyz\sum bc$. Is $abc\sum bc \geqq xyz\sum yz$?

I. *Solution by Simeon Reich, Israel Institute of Technology, Haifa.* Without loss of generality, we can assume that $a \geqq b \geqq c$. If x, y, and z are all positive, the first inequality is equivalent to $1/x + 1/y + 1/z \geqq 1/a + 1/b + 1/c$ which follows from the obvious inequalities $1/y + 1/z \geqq 2/a$, $1/z + 1/x \geqq 2/b$, and $1/x + 1/y \geqq 2/c$. Note that in this case the inequality is equivalent to $r_a + r_b + r_c \geqq h_a + h_b + h_c$, where h_a, h_b, h_c are the altitudes and r_a, r_b, r_c are the exradii of the triangle with sides a, b, c.

In general, y and z are always positive by assumption. If $x = 0$, the inequality is obvious and if x is negative, the inequality is equivalent to $1/x + 1/y + 1/z \leqq 1/a + 1/b + 1/c$, which is true because $1/y \leqq 1/c$, $1/z \leqq 1/b$, and $1/x < 0 < 1/a$. Thus the first inequality has been established.

As for the second inequality, it is obviously true if $x = 0$. If x is negative, it may hold (take $a = 3$ and $b = c = 1$) and it may not hold (take $a = 5$ and $b = c = 1$). Suppose that x, y, and z are all positive. Then a, b, and c are the sides of a triangle with altitudes h_a, h_b, h_c, exradii r_a, r_b, r_c, inradius r, circumradius R, and area S. Since

$$2S = ah_a = bh_b = ch_c = xr_a = yr_b = zr_c = 2(rr_ar_br_c)^{\frac{1}{2}} = (2Rh_ah_bh_c)^{\frac{1}{2}},$$

and since $r_a + r_b + r_c = 4R + r$, it is readily seen that the inequality is equivalent to $h_a + h_b + h_c \geqq 4r^2(4R + r)/R^2$. This follows from the fact that $R \geqq 2r$ and the

known inequality $h_a + h_b + h_c \geqq 2r(5R - r)/R$ (O. Bottema et al., *Geometric Inequalities*, Groningen, 1968, p. 63).

II. *Comment by Michael Goldberg, Washington, D.C.* We consider only the case of positive x, y, z. From the given relations, it follows that $x+y+z = a+b+c$ and that $a = \frac{1}{2}(y+z)$, $b = \frac{1}{2}(x+z)$, $c = \frac{1}{2}(x+y)$. Thus a, b, c are the arithmetic means of x, y, z taken in pairs. Since the means have the same sum as the original x, y, z, and since they are more nearly equal to each other, it follows that $abc \geqq xyz$ and $\sum bc \geqq \sum yz$. By multiplying these, we obtain the second inequality of the problem.

To show the first inequality, consider a, b, c to be the edges of a rectangular parallelepiped and x, y, z to be the edges of another. The volumes of the parallelepipeds are abc and xyz respectively and the surface areas are $2\sum bc$ and $2\sum yz$ respectively. If we take the ratio of volume to surface area, then $abc/\sum bc \geqq xyz/\sum yz$ since the first parallelepiped is more nearly a cube. This yields the first inequality.

Also solved by L. Carlitz, Frederick Carty, R. J. Dickson, Ralph Garfield, M. G. Greening (Australia), Robert Heller, Harry Lass, A. J. Patsche, David Spear, L. E. Ward, Sr., the proposer, and one solver whose solution was unsigned.

The proposer remarks that the first inequality in the case of positive x, y, z can be found in S. Barnard and J. M. Child, *Higher Algebra* (1936), p. 217. The proof there is similar to I above.

A Generalization of Napoleon's Theorem

E 2285 [1971, 297]. *Proposed by A. W. Walker, Toronto, Canada*

If X, Y, Z are similarly situated points of directly similar coplanar triangles DCB, CEA, BAF annexed to any triangle ABC, then triangle XYZ is directly similar to the annexed triangles.

I. *Solution by Leonard Goldstone, N. Y. State Department of Transportation.* We use the notation of David Merriell's paper, *An application of quasigroups to geometry*, this MONTHLY 77(1970), 44–46. Let the two operations be $\triangle$ for the given species triangle and $\circ$ for the homologous points. That is, if P and Q are any two points and if $R = P \triangle Q$, then triangle PQR is directly similar to triangle DCB and if $S = P \circ Q$, then triangle PQS is directly similar to triangle DCX. By hypothesis, we have that $B = D \triangle C$, $A = C \triangle E$, $F = B \triangle A$ and that $X = D \circ C$, $Y = C \circ E$, $Z = B \circ A$. Consider now $X \triangle Y = (D \circ C) \triangle (C \circ E)$; by Equation (6) of the reference, this is equal to $(D \triangle C) \circ (C \triangle E) = B \circ A = Z$, which proves the assertion.

II. *Solution by M. G. Greening, University of New South Wales, Australia.* Let α, β, δ be direct similitudes such that $\alpha(DCB) = CEA$, $\beta(CEA) = BAF$, $\delta(D) = X$, $\delta(C) = Y$, where DCB denotes triangle DCB, etc. Then α, β, and δ are uniquely determined. Now $Y = \alpha(X)$ since X and Y are similarly situated points of triangles DCB and CEA respectively. Then $Y = \delta(C) = \delta\alpha(D)$ and $Y = \alpha(X) = \alpha\delta(D)$. Now it is

readily seen that $(\alpha\delta)^{-1}(\delta\alpha)$ is a translation; since it has the fixed point D, it must be the identity, so that $\alpha\delta = \delta\alpha$. Similarly $\alpha\beta(C) = \alpha(B) = A$ and $\beta\alpha(C) = \beta(E) = A$, so that $\alpha\beta = \beta\alpha$. But two direct similitudes commute if and only if they have the same invariant point, so that α, β, δ all have the same invariant point and hence $\beta\delta = \delta\beta$. Consequently $\delta(B) = \delta\beta(C) = \beta\delta(C) = \beta(Y) = Z$, and $\delta(DCB) = XYZ$.

III. *Solution by Simeon Reich, Israel Institute of Technology, Haifa.* We identify in the usual way the point A with the complex number a, etc. There exist real scalars m_1, m_2, m_3 with $m_1 + m_2 + m_3 = 1$ such that $x = m_1 d + m_2 c + m_3 b$, $y = m_1 c + m_2 e + m_3 a$, $z = m_1 b + m_2 a + m_3 f$. Since triangles MNP and QRS are directly similar if and only if

$$\begin{vmatrix} m & n & p \\ q & r & s \\ 1 & 1 & 1 \end{vmatrix} = 0,$$

it follows that

$$\begin{vmatrix} x & y & z \\ d & c & b \\ 1 & 1 & 1 \end{vmatrix} = m_1 \begin{vmatrix} d & c & b \\ d & c & b \\ 1 & 1 & 1 \end{vmatrix} + m_2 \begin{vmatrix} c & e & a \\ d & c & b \\ 1 & 1 & 1 \end{vmatrix} + m_3 \begin{vmatrix} b & a & f \\ d & c & b \\ 1 & 1 & 1 \end{vmatrix} = 0 + 0 + 0 = 0,$$

as required.

IV. *Solution by the proposer.* The result can be proved by two applications of the following theorem: *Given two directly similar coplanar triangles $X_1Y_1Z_1$ and $X_2Y_2Z_2$, and points X_3, Y_3, Z_3 such that the ratios of directed segments satisfy $X_1X_3/X_1X_2 = Y_1Y_3/Y_1Y_2 = Z_1Z_3/Z_1Z_2$, then triangle $X_3Y_3Z_3$ is similar to the given triangles.* [*Editor's comment*: Note that X_1X_3/X_1X_2 will be negative if X_1 lies between X_2 and X_3 and positive otherwise. The points X_1, X_2, X_3 are by assumption collinear. The same holds for the Y's and Z's.] This theorem is a special case of Theorem 3.3.15 in H. Eves, *A Survey of Geometry*, Vol. I, Boston, 1963, p. 140. To prove the result, let I be the intersection of DX and BC, J the intersection of CY and AE, and K the intersection of BZ and FA. Then we have the directed ratios $BI/BC = AJ/AE = FK/FA$ and $DX/DI = CY/CJ = BZ/BK$. Applying the theorem to the similar triangles BAF and CEA and the points I, J, K, we see that triangles DCB and IJK are directly similar; applying the theorem to triangles DCB and IJK and points X, Y, Z, we have the result.

Also solved by R. J. Dickson, Jordi Dou (Spain), O. P. Lossers (Netherlands), and J. G. Mauldon.

A Decreasing Sequence

E 2286 [1971, 297]. *Proposed by E. T. H. Wang, University of British Columbia*

For each positive integer n, define $f(n)$ as $f(n) = (n!)^{1/n}$. Prove or disprove that the sequence

$$\left\{\frac{f(n+1)}{f(n)}\right\}_{n=1}^{\infty}$$

is monotonically decreasing.

I. *Solution by Michael Schulz, The Aerospace Corporation, El Segundo, California.* It is easy to verify by direct evaluation that the sequence decreases monotonically for $n = 1,2,3,4$. It is necessary to prove in general that

$$\frac{f(n+2)}{f(n+1)} \Big/ \frac{f(n+1)}{f(n)} < 1.$$

To achieve this, define the function

$$R(n) \equiv \left\{\frac{f(n+2)f(n)}{[f(n+1)]^2}\right\}^{n(n+1)(n+2)}$$

and assume $R(n) < 1$ as the induction hypothesis. It follows that

$$R(n+1) = R(n)\left[\frac{n^2+4n+3}{n^2+4n+4}\right]^{(n+1)(n+2)} < 1$$

for every positive integer n.

II. *Solution by W. C. Taylor and B. H. Rodin, Aberdeen Proving Ground, Maryland.* We show equivalently that

$$F_n \equiv \frac{f(n+1)}{f(n)} \Big/ \frac{f(n)}{f(n-1)} < 1 \quad \text{for} \quad n = 2, 3, \cdots.$$

Raising F_n to a power we find

$$(F_n)^{n(n+1)/2} = [(n-1)!]^{1/(n-1)} \frac{1}{n} \left(\frac{n+1}{n}\right)^{n/2}.$$

Since the geometric mean is less than the arithmetic mean,

$$[(n-1)!]^{1/(n-1)} < \frac{(n-1)+(n-2)+\cdots+3+2+1}{n} = \frac{n(n-1)}{2n} < \frac{n}{2}.$$

Therefore,

$$(F_n)^{n(n+1)/2} < \tfrac{1}{2}\left(1+\frac{1}{n}\right)^{n/2} < \tfrac{1}{2}e^{\frac{1}{2}} < 1.$$

Thus $F_n < 1$ and the sequence $\{f(n+1)/f(n)\}_{n=1}^{\infty}$ is monotonically decreasing.

III. *Remark by L. Carlitz, Duke University.* Minc and Sathre [Proc. Edinburgh Math. Soc. (2), 14 (1964–5), 41–46] have proved that

$$\{f(n)\}, \quad \left\{\frac{n}{f(n)}\right\}, \quad \left\{\frac{nf(n+1)}{f(n)}\right\}$$

are strictly increasing, and that

$$1 < \frac{f(n+1)}{f(n)} < \frac{n+1}{n}.$$

Also solved by seventy-one other readers.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before May 31, 1972. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5838. *Proposed by R. B. Eggleton, University of Calgary*

Let $N(g)$ denote the number of isomorphism classes of abelian groups of order g. The equation $N(x) = n$ is solvable for $1 \leqq n \leqq 12$, and for infinitely many other natural numbers n, but there is no solution when $n = 13$. Show that there are infinitely many natural numbers n for which there is no solution.

5839. *Proposed by A. D. Ziebur, State University of New York, Binghamton*

The equation $\pi(x, y) = x^y$ defines a function from $R^+ \times R$ to R^+ (R is the set of real numbers, R^+ the set of positive reals) such that $\pi(x, n) = x^n$ when n is an integer, and $\pi(x, yz) = \pi(\pi(x, y), z)$. Is the power function the only function with these properties?

5840.* *Proposed by Maury Horowitz and Nick Metas, Queens College, and Gerald Leibowitz, University of Connecticut*

Can one construct a real-valued function f whose domain is an open set U in R^2 such that f has all partial derivatives of all orders at every point of U, yet there is some point of U at which f is not continuous?

5841. *Proposed by L.-S. Hahn, University of New Mexico*

Is there a (complex) continuous measure (i.e., $\mu(E) = 0$ if E is countable) on the real line, whose Fourier-Stieltjes transform has modulus 1 everywhere on the real line?

SOLUTIONS OF ADVANCED PROBLEMS

Fixed Points of Minkowski's Singular Monotone Function

5768 [1970, 1115]. *Proposed by Peter Flor, University of Vienna, Austria*

Minkowski's singular monotone function $M(x)$ is defined as follows: $M(0) = 0$, $M(1) = 1$; if $a = p/q$ and $b = p'/q'$ are rational numbers such that $p'q - pq' = 1$, and if $c = (p + p')/(q + q')$, then $M(c) = \frac{1}{2}[M(a) + M(b)]$. This defines $M(x)$ for every rational $x \in [0, 1]$; the function can then be extended by continuity to all of $[0, 1]$. Obviously 0, $\frac{1}{2}$, and 1 are fixed points of $M(x)$. Prove that there are exactly two further fixed points, $d_1 = 0.4203723\cdots$ and $d_2 = 1 - d_1$. Decide whether they are rational.

Solution by L. E. Mattics, University of South Alabama. We shall show that d_1 and d_2 are the only other fixed points and that they are irrational. Since $M(1-x) = 1 - M(x)$ we have only to study $M(x)$ for $x \in [0, \frac{1}{2}]$. Using the monotonicity of $M(x)$ we first note

(I) *If p/q and p'/q' are rationals in $[0,1]$ with $p'/q' > p/q$ and $p'q - pq' = 1$, then $p/q > M(p'/q')$ implies $x > M(x)$ and also $M(p/q) > p'/q'$ implies $M(x) > x$, for all $x \in [p/q, p'/q']$.*

If we now let F_7 be the sequence of Farey fractions of order 7 between 0 and $\frac{1}{2}$ inclusive, we note that for $x \in F_7$, $M(x) > x$ if $\frac{1}{2} > x \geqq 3/7$ and $x > M(x)$ if $2/5 \geqq x > 0$. Now if $n \geqq 4$, $1/(n + 1) > M(1/n) = 1/2^{n-1}$, so $x > M(x)$ for $x \in (0, 1/4]$ by (I). On the intervals $[1/4, 2/7]$, $[2/7, 1/3]$, $[1/3, 3/8]$, $[3/8, 5/13]$, $[5/13, 2/5]$, $[2/5, 7/17]$, $[7/17, 5/12]$, (I) applies to show that $x > M(x)$ on $(0, 5/12]$. Similarly, applying (I) to the intervals $[8/19, 35/83]$, $[35/83, 19/45]$, $[19/45, 11/26]$, $[11/26, 3/7]$, $[3/7, 10/23]$, $[10/23, 4/9]$ shows that $M(x) > x$ on $[8/19, 4/9]$. Since

$$M\left(\frac{3 + n}{7 + 2n}\right) - \frac{3 + (n + 1)}{7 + 2(n + 1)} = \frac{2^{n+3} - (2n + 9)}{2^{n+4}(9 + 2n)} > 0$$

for $n \geqq 1$, we have by (I) that $M(x) > x$ on $[4/9, 1/2)$; so $M(x) > x$ on $[8/19, 1/2)$.

We now define the sequences $\{s_i\}$, $\{t_i\}$ by $s_1 = 5/12$ and $t_1 = 8/19$ and if $s_i = p/q$ and $t_i = p'/q'$, then $s_{i+1} = (p + p')/(q + q')$ if $(p + p')/(q + q') \geqq M((p + p')/(q + q'))$, and $s_{i+1} = s_i$ otherwise; similarly $t_{i+1} = (p + p')/(q + q')$ if $M((p + p')/(q + q')) \geqq (p + p')/(q + q')$ and $t_{i+1} = t_i$ otherwise. Now $\{s_i\}$ and $\{t_i\}$ converge to a common limit, say d_1, and $M(d_1) = d_1$.

If d_1 were rational then by the theory of Farey fractions and the definitions of $\{s_i\}$ and $\{t_i\}$ there would exist n such that $s_j = t_j = d_1$, and thus $s_j - M(s_j) = t_j - M(t_j) = 0$ for all $j \geqq n$. To show that d_1 is irrational and that there are no other fixed points between 5/12 and 8/19 we prove

(II) *For any i, $s_i - M(s_i) > 1/2^{5+i}$ and $M(t_i) - t_i > 1/2^{5+i}$.*

We do only the first part. Note $5/12 - M(5/12) > 1/64$, $13/31 - M(13/31) > 1/128$,

$21/50 - M(21/50) > 1/256$, and also $s_3 = s_4$ and $t_1 = t_2 = t_3$. So, suppose $s_i - M(s_i) > 1/2^{5+i}$ for all i, $1 \leqq i \leqq n$ and $n \geqq 3$. If $s_n = s_{n+1}$ then the induction step is trivial. If $s_{n+1} > s_n$ then there is a smallest p, $0 \leqq p \leqq n-1$ such that $s_{n-p} = s_{n-(p+1)}$. Now

$$s_{n+1} - M(s_{n+1}) = s_{n+1} - s_{n-(p+1)} + (s_{n-(p+1)} - M(s_{n-(p+1)})) - \sum_{j=0}^{p+1} 1/2^{6+n-j}$$

$$> 1/2^{6+n-p-1} - (1/2^{6+n-p}) \sum_{j=0}^{p+1} 1/2^{j} = 1/2^{5+(n+1)}.$$

[Note that $M(t_1) - M(s_1) = 1/2^6$, and thus $M(s_{i+1}) - M(s_i) = 1/2^{6+i}$ for $n - p \leqq i \leqq n$.]

Hence (II) is proved by induction. We finish by noting that by (II) $s_n - M(s_{n+1}) = s_n - M(s_n) + M(s_n) - M(s_{n+1}) \geqq s_n - M(s_n) - 1/2^{6+n} > 0$, so by (I) $x > M(x)$ on $[5/12, d_1)$. Similarly $M(x) > x$ on $(d_1, 8/19]$.

The Clone of Ternary Majority Functions, I

5771 [1971, 83]. *Proposed by G. M. Bergman, Bedford College, London, England*

On the set $\{0,1\}$, let $(\ , \ , \ )$ designate the ternary "majority vote" operation defined by:

$$(a,b,c) = \begin{cases} 0 \text{ if at least two of } a,\ b,\ c \text{ are } 0 \\ 1 \text{ if at least two of } a,\ b,\ c \text{ are } 1. \end{cases}$$

Consider the clone of operations this generates—e.g., this contains the 4-ary operation $x(a,b,c,d) = (a,b,(c,d,a))$, the 9-ary operation $y(a,\cdots,i) = ((a,b,c),(d,e,f),(g,h,i))$, etc.

Prove that an n-ary operation $f\colon \{0,1\}^n \to \{0,1\}$ lies in this clone if and only if (1) $a_i \leqq b_i$ $(i = 1,\cdots,n)$ implies $f(a_1,\cdots,a_n) \leqq f(b_1,\cdots,b_n)$ and (2) $f(1-a_1,\cdots,1-a_n) = 1 - f(a_1,\cdots,a_n)$.

I. *Solution by Joel Spencer, RAND Corporation, Santa Monica, Cal.* By an obvious induction, if f is in the clone then f satisfies (1) and (2). Let $S \subseteq \{1,2,\cdots,n\}$. Write $f(S) = f(x_1,\cdots,x_n)$ where $x_i = 1$ if and only if $i \in S$. Set $F = \{S : f(S) = 1\}$. Then F satisfies (1′) $S \in F$ if and only if $S^c \notin F$; and (2′) $S \in F$ and $S \subseteq T$ imply $T \in F$. If any $\{i\} \in F$, then $F = \{S : i \in S\}$, whence $f(a_1,\cdots,a_n) = a_i$ is in the clone. If $\{i\} \notin F$, order $F = \{S_1,\cdots,S_m\}$ in any manner. Let f_0 be any member of the clone. Having defined f_{k-1}, if $S_k = \{x_1,\cdots,x_j\}$ set

$$f_k = (x_1, f_{k-1}, (x_2, f_{k-1}, (\cdots (x_{j-2}, f_{k-1}, (x_{j-1}, x_j, f_{k-1}) \cdots).$$

By induction, f_k is in the clone. Setting $F_k = \{S : f_k(S) = 1\}$,

$$F_k = \{T : S_k \subseteq T\} \cup \{T : T \in F_{k-1},\ S_k^c \nsupseteq T\}.$$

Each $S_i \in F_i$. By (1′) and (2′) we cannot have $S_i^c \supseteq S_j$. Therefore, each $S_i \in F_m$. Thus $F \subseteq F_m$ and since both F and F_m satisfy (1′) we must have $F = F_m$, and so $f = f_m$ is in the clone.

This representation of f may have "length" in excess of 2^{2^n}. It would be interesting to see if the minimal length of a representation of f could be substantially reduced.

II. *Solution (Abstract by the Editors) by Frank R. Bernhart, University of Kansas.* Each function satisfying (1) and (2) is constructible by ternary majority from the functions $g_i(a_1, \cdots, a_n) = a_i$ in the following way. A *geometry* is defined as a family of subsets (called lines) of a finite set X satisfying: (G1) No line properly contains another line, and (G2) no two lines are parallel (the intersection of each pair is nonempty), and (G3) in each dichotomy of X into two sets, at least one of the sets contains a line.

Let (without loss of generality) $X = \{x_1, \cdots, x_n\}$, n odd. Define $f(A)$, $A \subseteq X$, to be the value of f when each x_i is replaced by 1 if $x_i \in A$, by 0 otherwise. A one-to-one correspondence is established between functions satisfying (1) and (2) and geometries: Given f, let $G = G(f)$ be the collection of minimal subsets A of X so that $f(A) = 1$. Given geometry G, let $f = f_G$ be the function such that $f(A) = 1$ if and only if A contains a line of G. Then if $G = G(f), f = f_G$. The result is then established by induction on the number of small lines (those with less than $n/2$ members).

The following additional questions offer some challenge: (1) To enumerate the geometries (or functions) for a given n, either including or excluding isomorphic types. (2) Define the majority functions M_k for $k = 1, 2, \cdots$ on the set $\{+1, -1\}$ as symmetric functions such that $M_k(x, x, \cdots, x) = x$, and $M_{k+1}(x_1, x_2, \cdots, x_{2k-1}, y, -y) = M_k(x_1, x_2, \cdots, x_{2k-1})$. Describe the clone generated by each M_k. (3) Let P_n^r denote the homogeneous function of degree r in x_i, $i = 1, 2, \cdots, n$, with unit coefficients over the modulus 3 system $\{+1, 0, -1\}$. Then we find $M_1 = P_1^1$, $M_2 = P_3^3 - P_3^1$, and $M_3 = P_5^3$. Show that M_k has a representation in the following form, and find the coefficients a_{ki}:

$$M_k = \sum_{i=1}^{k} a_{ki} P_{2k-1}^{2i-1} \pmod 3.$$

Also solved by the proposer.

The Clone of Ternary Majority Functions, II

5772 [1971, 83]. *Proposed by G. M. Bergman, Bedford College, London, England*

Can every function f with properties (1) and (2) listed in the preceding problem be constructed as a "weighted vote" function? That is, given such an f, can we always find $\lambda_1, \cdots, \lambda_n \in [0, 1]$ summing to 1, and having no subset summing to exactly $\frac{1}{2}$, such that

$$f(a_1, \cdots, a_n) = \begin{cases} 0 & \text{if} \quad \sum \lambda_i a_i < \frac{1}{2} \\ 1 & \text{if} \quad \sum \lambda_i a_i > \frac{1}{2} \end{cases} ?$$

Solution by Joel Spencer, RAND Corporation, Santa Monica, Cal. No. Say

$$f(a_1,\cdots,a_9) = ((a_1,a_2,a_3),(a_4,a_5,a_6),(a_7,a_8,a_9)).$$

Assume weights $\lambda_1,\cdots,\lambda_9$ could be assigned. Set $A = \lambda_1 + \lambda_2 + \lambda_3$, $B = \lambda_4 + \lambda_5 + \lambda_6$, $C = \lambda_7 + \lambda_8 + \lambda_9$. By symmetry we may assume $A \leqq B \leqq C$, $\lambda_1 \leqq \lambda_2 \leqq \lambda_3$, and $\lambda_4 \leqq \lambda_5 \leqq \lambda_6$. Then setting $a_i = 1$, if $i = 1$, 2, 4 or 5, and setting $a_i = 0$ elsewhere, we have $f(a_1,\cdots,a_9) = 1$, but $\sum \lambda_i a_i \leqq 4/9$.

Also solved by D. R. Anderson, J. M. Reiner, and the proposer.

Finite Cyclic Groups

5774 [1971, 84]. *Proposed by J. C. Owings, Jr., University of Maryland*

Let G be a finite group and suppose, for all $d \geqq 1$, that G has at most d elements of order d. Prove G is cyclic.

Solution by S. J. Tillman, Wilkes College. Suppose $|G| = n$, and that $d \mid n$. Let A_d be the set of all elements of G whose exact order is d. Suppose $a \in A_d$. Then $e, a^1,\cdots,a^{d-1}$ all satisfy $x^d = e$, where e is the multiplicative identity of G. By hypothesis these must be the only elements of G which do so. Hence either $|A_d| = 0$, or $|A_d| = \phi(d)$, where ϕ is the Euler ϕ-function. Clearly $A_{d_2} \cap A_{d_1} = \emptyset$ if $d_1 \neq d_2$ and $G = \cup_{d|n} A_d$. Hence

$$n = \sum_{d|n} |A_d| \leqq \sum_{d|n} \phi(n) = n.$$

Thus $|A_d| = \phi(d)$, so in particular $|A_n| = \phi(n)$, so G has an element of order n, so is cyclic.

Editorial Notes. (1) D. M. Bloom points out that the result is Theorem 5.7.6, p. 118 in W. R. Scott, *Group Theory*, and Lindsay Childs finds the problem as 11.18 on page 95 of Fraleigh, *A First Course in Abstract Algebra.*

(2) J. H. E. Cohn, in a forthcoming paper in the Proc. A. M. S., *A condition for a finite group to be cyclic*, proves the following generalizations:

(a) G is cyclic if for every prime power $q = p^k$, the equation $x^q =$ identity has at most $p^{k+1}-1$ solutions.

(b) G is cyclic if for every prime power $q = p^k$, there are at most $p^{k-1}(p^2-1) - 1$ elements of order q precisely.

Also solved by the proposer and thirty-two other contributors.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 3

CONTENTS

(Continued on inside cover)

MARCH 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the Monthly. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to Harley Flanders, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

Published by the Association at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

THE FORMATION AND DECAY OF SHOCK WAVES

PETER D. LAX, Courant Institute, New York University

1. Introduction. The theory of propagation of shock waves is one of a small class of mathematical topics whose basic problems are easy to explain but hard to resolve. This article is a brief introduction to the subject: we shall describe the origin of the governing equations, some of the striking phenomena, and a few of the mathematical tools used to analyse them.

2. What is a conservation law? A conservation law asserts that the change in the total amount of a physical entity contained in any region G of space is due to the **flux** of that entity across the boundary of G. In particular, the rate of change is

$$\frac{d}{dt}\int_G u\,dx = -\int_{\partial G} f\cdot n\,dS, \tag{2.1}$$

where u measures the **density** of the physical entity under discussion, and the vector f describes its flux; n is the outward normal to the boundary ∂G of G. If u and f are differentiable functions, we can, on the left, perform the differentiation under the integral sign and on the right apply the divergence theorem. We obtain

$$\int_G \{u_t + \operatorname{div} f\}\,dx = 0.$$

This relation is assumed to be valid for every domain G. Letting G shrink to a point and dividing by the volume of G we get the differential form of the conservation law:

$$u_t + \operatorname{div} f = 0. \tag{2.2}$$

To complete the theory we need some law relating f to u. E.g., Newton's law of cooling asserts that the flux of heat is proportional to the negative gradient of u, where u is temperature; in this case $f = -h$ grad u, h positive, so (2.2) becomes

$$u_t - h\Delta u = 0, \qquad \Delta = \text{div grad}.$$

In this example f depends on the derivatives of u; in what follows *we assume that f depends on u alone.* More precisely, we shall be looking at systems of conservation laws

$$u_t^j + \operatorname{div} f^j = 0, \qquad j = 1,\cdots,n, \tag{2.3}$$

where each f^j is a function of all the $u^1,\cdots,u^n$, and a nonlinear function at that.

Peter Lax received his Ph.D. at New York University under K. Friedrichs and has spent most of his academic career at New York University, where he is presently a professor. He is a frequent summer visitor at Stanford and the Los Alamos Scientific Lab. His research contributions in partial differential equations, linear and non-linear problems of mathematical physics, computing, and functional analysis have had a profound impact. He was a Fullbright lecturer in 1958, he is a Vice-President of the AMS, he is an elected member of the National Academy of Sciences, he was an AMS Gibbs lecturer, and he received an MAA Lester Ford Award. He is co-author with R. Phillips of *Scattering Theory* (Academic Press, 1967). *Editor.*

Many equations of mathematical physics are of this form, in particular, those governing the flow of a nonviscous, compressible fluid.

We shall concern ourselves with the **initial value problem** for systems of form (2.3); that is, given the value of each u^j at $t = 0$ as function of x, determine u^j as function of x and t for all $t > 0$.

3. The theory of a single nonlinear conservation law. In this section we shall study conservation laws for a single quantity u dependent on only one space variable x; in this case f has only one component:

$$u_t + f_x = 0, \tag{3.1}$$

where f is some nonlinear function of u. Denoting

$$\frac{df}{du} = a(u) \tag{3.2}$$

we can write (3.1) in the form

$$u_t + a(u)u_x = 0 \tag{3.3}$$

which asserts that u is constant along trajectories $x = x(t)$ which propagate with speed a:

$$\frac{dx}{dt} = a. \tag{3.4}$$

For this reason a is called the **signal speed**; the trajectories, satisfying (3.4), are called **characteristics.** Note that if f is a nonlinear function of u, both signal speed and characteristics depend on the solution u.

The constancy of u along characteristics combined with (3.4) shows that the characteristics propagate with constant speed; so they are straight lines. This leads to the following geometric solution of the initial value problem

$$u(x, 0) = u_0(x).$$

Draw straight lines issuing from points y of the x-axis, with slope $1/u_0(y)$ (see Fig. 1).

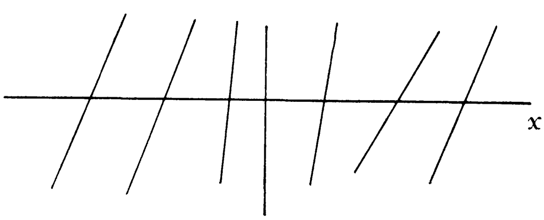


FIG. 1

As we shall show, if u_0 is a C^1 function, these lines simply cover a neighborhood of the x-axis; since the value of u along the line issuing from the point y is $u_0(y)$, $u(x,t)$ is uniquely determined near the x-axis.

An analytical form of this construction goes like this (see Fig. 2)

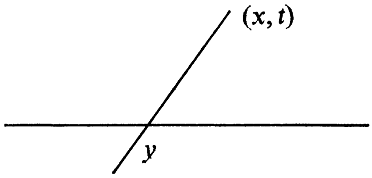


FIG. 2

Let (x,t) be any point, y the intersection of the characteristic through x, t with the x-axis. Then $u = u(x,t)$ satisfies

$$u = u_0(y),\ y = x - t\,a(u). \tag{3.5}$$

Assume u_0 differentiable; then, according to the implicit function theorem, (3.5) can be solved for u as a differentiable function of x and t for t small enough, and

$$u_t = -\frac{u_0' a}{1 + u_0' a_u t} \qquad u_x = \frac{u_0'}{1 + u_0' a_u t}. \tag{3.6}$$

Substituting (3.6) into (3.3) we see immediately that u defined by (3.5) satisfies (3.3).

Let's assume that equation (3.3) is **genuinely nonlinear**, i.e., that $a_u \neq 0$ for all u, say

$$a_u > 0. \tag{3.7}$$

Then if u_0' is $\geqq 0$ for all x, u_t and u_x as given by formulas (3.6) remain bounded for all $t > 0$; on the other hand, if u_0' is < 0 at some point, both u_t and u_x tend to ∞ as $1 + u_0'\, a_u(u_0)t$ approaches zero. Both these facts can be deduced from the geometric form of the solution contained in Figure 1:

In the first case, when $u_0(x)$ is an increasing function of x, the characteristics issuing from the x-axis diverge in the positive t direction, so that the characteristics simply cover the whole half-plane $t > 0$. In the second case there are two points y_1 and y_2 such that $y_1 < y_2$, and $u_1 = u_0(y_1) > u_0(y_2) = u_2$; then by (3.7) also $a_1 = a(u_1) > a(u_2) = a_2$ so that the characteristics issuing from these points intersect at time

$$t = \frac{y_2 - y_1}{a_1 - a_2}.$$

At the point of intersection, u has to take on the value u_1 and u_2 both, an impossibility (see Fig. 3).

Both the geometric and the analytic argument prove beyond the shadow of a

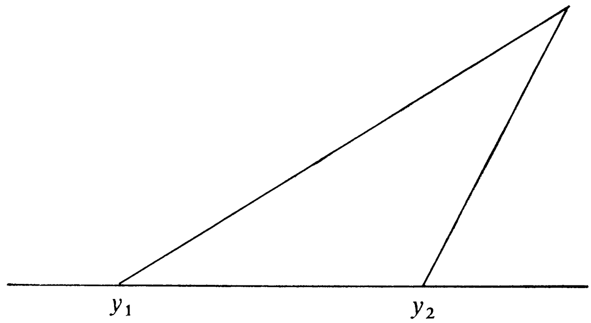


FIG. 3

doubt that if the initial value u_0 is not an increasing function of x then *no continuous function* $u(x,t)$ *exists for all* $t > 0$ *with initial value* u_0 *which solves equation* (3.3) *in the ordinary sense*!

What happens after continuous solutions cease to exist? After all, the world does not come to an end. For an answer, we turn to experiments with compressible fluids: these clearly show the appearance of discontinuities in solutions. We begin our study of discontinuous solutions with the simplest kind, those satisfying (3.1) in the ordinary sense on each side of a smooth curve $x = y(t)$ across which u is discontinuous. We shall denote by u_l and u_r the values of u on the left and right sides respectively of $x = y(t)$. Choose a and b so that the curve y intersects the interval $a \leqq x \leqq b$ at time t (see Fig. 4).

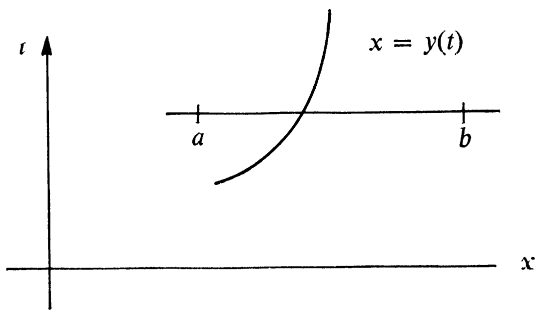


FIG. 4

Denoting by $I(t)$ the quantity $I(t) = \int_a^b u(x,t)\,dx = \int_a^y + \int_y^b$, we have

$$\frac{dI}{dt} = \int_a^y u_t dx + u_l s + \int_y^b u\,dx - u_r s, \tag{3.8}$$

where we have used the abbreviation

$$s = \frac{dy}{dt} \tag{3.9}$$

for the speed with which the discontinuity propagates. Since on either side of the discontinuity (3.1) is satisfied we may set $u_t = -f_x$ in the integrals in (3.8); after carrying out the integration we obtain $dI/dt = f_a - f_l + u_l s - f_b + f_r - u_r s$; here we have used the handy abbreviations

$$f(u_l) = f_l, \qquad f(u_r) = f_r$$
$$f(u(a)) = f_a, \qquad f(u(b)) = f_b.$$

The conservation law asserts that $dI/dt = f_a - f_b$. Combining this with the above relation we deduce the **jump condition**

$$s[u] = [f], \tag{3.10}$$

where $[u] = u_r - u_l$ and $[f] = f_r - f_l$ denote the jump in u and in f across y.

We show now in an example that previously unsolvable initial value problems can be solved for all t with the aid of discontinuous solutions. Take

$$f(u) = \tfrac{1}{2}u^2, \tag{3.11}$$

$$u_0(x) = \begin{cases} 1 & \text{for } x \leqq 0 \\ 1 - x & \text{for } 0 \leqq x \leqq 1 \\ 0 & \text{for } 1 \leqq x. \end{cases}$$

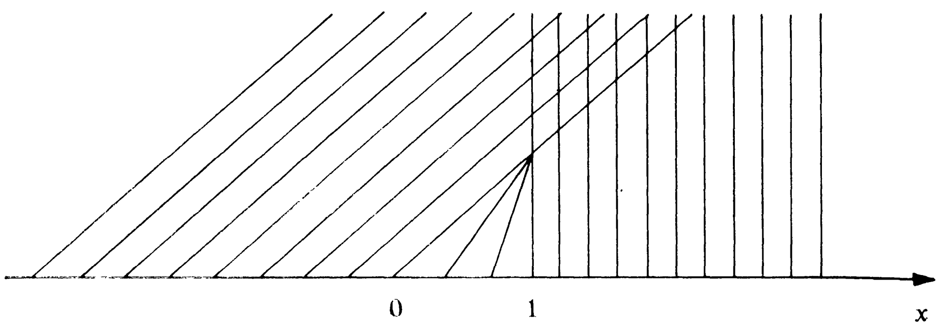


FIG. 5

The geometric solution is single valued for $t \leqq 1$ but double valued thereafter (see Fig. 5). Now we define for $t \geqq 1$

$$u(x,t) = \begin{cases} 1 \text{ for } x < (1+t)/2 \\ 0 \text{ for } (1+t)/2 < x. \end{cases}$$

The discontinuity starts at (1,1); it separates the state $u_l = 1$ on the left from the state $u_r = 0$ on the right; the speed of propagation was chosen according to the jump condition (3.10), with $f(u) = \frac{1}{2}u^2$:

$$s = \frac{0 - \frac{1}{2}}{0 - 1} = \frac{1}{2}.$$

Introducing generalized solutions makes it possible to solve intial value problems which could not be solved within the class of genuine solutions. At the same time there is the danger that the enlarged class of solutions is so large that there are several generalized solutions with the same initial data. The following example shows that this anxiety is well founded:

$$u_0(x) = \begin{cases} 0 \text{ for } x < 0 \\ 1 \text{ for } 0 < x. \end{cases}$$

The geometric solution

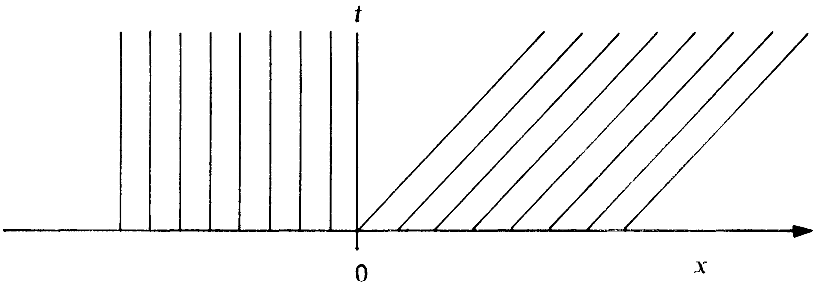


FIG. 6

is single valued for $t > 0$ (see Fig. 6) but does not determine the value of u in the wedge $0 < x < t$. We could fill this gap in the fashion of the previous example and set

$$u(x,t) = \begin{cases} 0 \text{ for } x < t/2 \\ 1 \text{ for } t/2 < x. \end{cases} \tag{3.12}$$

The speed of propagation was so chosen that the jump condition (3.10) is satisfied. On the other hand the function

$$u(x,t) = x/t, \qquad 0 \leqq x \leqq t \tag{3.12'}$$

satisfies the differential equation (3.3) with $a(u) = u$, and joins continuously the rest of the solution determined geometrically. Clearly only one of these solutions can have physical meaning; the question is which?

We reject the discontinuous solution (3.12) for failure to satisfy the following criterion:

The characteristics starting on either side of the discontinuity curve when continued in the direction of positive t intersect the line of discontinuity. This will be the case if

$$a(u_l) > s > a(u_r). \tag{3.13}$$

Under condition (3.7) for a this means that

(3.14) $$u_l > u_r.$$

Clearly this condition is violated in the solution given by (3.12).

The analysis at the beginning of this section shows that signals propagate along characteristics. Condition (3.13) allows each point of the discontinuity to be reached by characteristics on both sides, so that the shock is influenced by the initial data of the solution; this constitutes one justification of Condition (3.13). Another justification can be based on characterising the physically meaningful solutions as limits, when u tends to zero, of the viscous equation

$$u_t + f(u)_x = \mu u_{xx}, \qquad \mu > 0.$$

Yet another justification can be based on the theory of entropy. We shall not go into this interesting matter any deeper here, but merely record the gratifying fact that when $a(u)$ is a monotonic function of u, condition (3.13) is restrictive enough to make the solution of the initial value problem unique, yet it is broad enough to allow the construction of a solution for all time $t > 0$, having as initial value any integrable function u_0. True, the concept of solution has to be generalized beyond simple discontinuities: a bounded measurable function $u(x,t)$ is said to satisfy the conservation law (3.1) in the sense of distributions, if for all continuously differentiable test functions $\phi(x,t)$, with support in $t > 0$,

(3.15) $$\int\int [\phi_t u + \phi_x f(u)]dxdt = 0.$$

It is easy to verify that for the previously considered class of piecewise continuous solutions condition (3.15) is equivalent with the jump condition (3.10).

For merely bounded, measurable solutions u_l and u_r in condition (3.13) have to be interpreted as follows:

$$u_l = \liminf_{y \to x,\ y<x} u(y,t),$$

$$u_r = \limsup_{y \to x,\ x<y} u(y,t).$$

For the main existence theorem we refer the reader to [**8**] and [**13**], and for uniqueness to [**1**], [**14**], and [**16**].

It turns out that when $a(u)$ is not monotonic, condition (3.13) is not sufficient to guarantee unique determination of solutions by their initial data. A replacement for this condition has been found by Oleinik; this condition, together with the existence and uniqueness theorem is described in [**15**]; other interesting discussions of this condition are contained in [**4**], [**6**], and [**16**].

4. The decay of solutions. Existence and uniqueness of solutions is not the

end but merely the beginning of a theory of differential equations. The really interesting questions concern the behavior of solutions.

Here we shall study the asymptotic behavior for large time of solutions of conservation laws of form (3.1) which satisfy condition (3.14); we assume that $a(u)$ is an *increasing* function of u.

As remarked in Section 3, any differentiable solution u is constant along characteristics

$$\frac{dx}{dt} = a(u) = f'(u). \tag{4.1}$$

Let $x_1(t)$ and $x_2(t)$ be a pair of characteristics, $0 \leqq t \leqq T$. Then there is a whole one-parameter family of characteristics connecting the points of the interval $[x_1(0), x_2(0)]$, $t = 0$ with points of the interval $[x_1(T), x_2(T)]$, $t = T$; since u is constant along these characteristics, $u(x,0)$ *on the first interval and* $u(x,T)$ *on the second interval are equivariant*, i.e., they take on the same values in the same order. Since equivariant functions have the same total increasing and decreasing variations, we conclude that *the total increasing and decreasing variations of a differentiable solution between any pair of characteristics are conserved.*

Denote by $D(t)$ the width of the strip bounded by x_1 and x_2:

$$D(t) = x_2(t) - x_1(t) > 0. \tag{4.2}$$

Differentiating (4.2) with respect to t and using (4.1), we get

$$\frac{d}{dt}D(t) = \frac{dx_2}{dt} - \frac{dx_1}{dt} = a(u_2) - a(u_1). \tag{4.3}$$

Integrating with respect to t we get

$$D(T) = D(0) + [a(u_2) - a(u_1)]T. \tag{4.4}$$

Suppose there is a shock y present in u between the characteristics x_1 and x_2 (see Fig. 7). Since according to condition (3.13) characteristics on either side of a shock run into the shock, there exist for any given time T two characteristics y_1 and y_2 which intersect the shock y at exactly time T. Assuming that there are no other shocks present we conclude that the increasing variation of u on $(x_1(t),\ y_1(t))$, as well as on $(x_2(t), y_2(t))$, is independent of t. According to condition (3.14), u decreases across shocks, so the increasing variation of u along $[z_1(T), x_2(T)]$ equals the sum of the increasing variations of u along $[x_1(0),\ y_1(0)]$ and along $[y_2(0), x_2(0)]$. This sum is in general less than the increasing variation of u along $[x_1(0), x_2(0)]$, therefore we conclude that if shocks are present, *the total increasing variation of u between two characteristics decreased with time.*

We give now a quantitative estimate of this decrease. Let I_0 be any interval of the x-axis; we subdivide it into subintervals $[y_{j-1}, y_j]$, $j = 1, \cdots, n$ in such a way

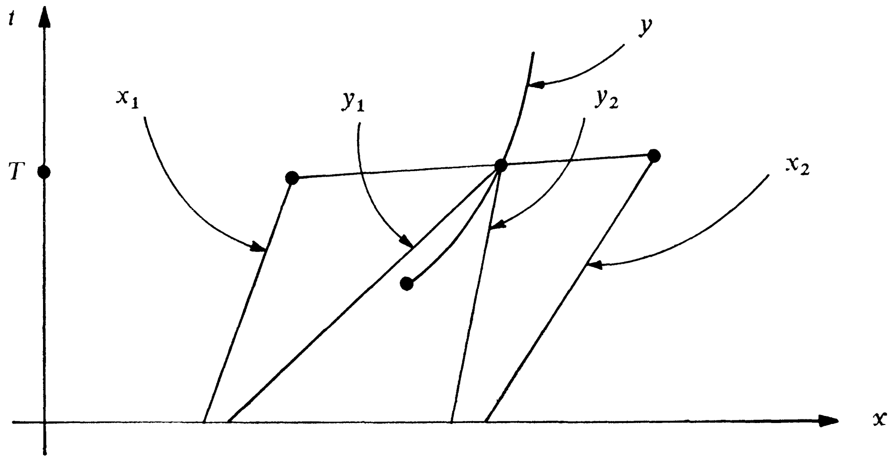


FIG. 7

that $u(x,0)$ is alternately increasing and decreasing on the intervals (we here assumed for simplicity that u_0 is piecewise monotonic). We denote by $y_j(t)$ the characteristic issuing from the jth point y_j, with the understanding that if $y_j(t)$ runs into a shock, $y_j(t)$ is continued as that shock.

It is easy to show that for any $t > 0$, $u(x,t)$ is alternately increasing and decreasing on the intervals $(y_{j-1}(t),\ y_j(t))$. Since a is an increasing function of u, and since according to (3.14) u decreases across shocks, the total increasing variation $A^+(T)$ of $a(u)$ across the interval $I(T) = [y_0(T),\ y_n(T)]$ is

$$\sum_{j\text{ odd}} a(u_j(T) - a(u_{j-1}(T)) = A^+(T), \tag{4.5}$$

where $u_{j-1}(T)$ denotes the value of u on the right edge of $y_{j-1}(T)$, $u_j(T)$ denotes the value of u on the left edge of $y_j(T)$; in case $y_{j-1}(T)$ and $y_j(T)$ are the same, the jth term in (4.5) is zero. Suppose $y_{j-1}(T)$ and $y_j(T)$ are shocks; then there exist characteristics $x_{j-1}(t)$ and $x_j(t)$ which start at $t = 0$ inside (y_{j-1}, y_j) and which at $t = T$ run into $y_{j-1}(T)$ and $y_j(T)$ respectively. The value of u along $x_j(t)$ is $u_j(T)$.

Denote $x_j(t) - x_{j-1}(t)$ by $D_j(t)$; according to (4.4)

$$D_j(T) = D_j(0) + [a(u_j) - a(u_{j-1})]T.$$

Summing over j odd and using (4.5) we get

$$\Sigma D_j(T) = \Sigma D_j(0) + A^+(T)T. \tag{4.6}$$

Since the intervals $[x_{j-1}(T), x_j(T)] = [y_{j-1}(T), y_j(T)]$ are disjoint and lie in $I(T)$, their total length cannot exceed the length $L(t)$ of $I(T)$; so we deduce from (4.6) that

$$A^+(T) \leqq \frac{L(T)}{T}, \tag{4.7}$$

where $A^+(T)$ is the total increasing variation of $a(u)$ along $I(T)$.

Let $u(x,t)$ be a solution of (3.1), possibly discontinuous, whose initial values are bounded, and zero outside a finite interval I_0. Since signals propagate with finite speed, for every t the solution $u(x,t)$ is zero outside some finite x-interval $I(t)$. Denote by $v(t)$ and $w(t)$ the values of u at the left and right endpoints of $I(t)$ respectively. Since the endpoints may lie on shocks, these values need not be zero, however it follows from (3.14) that

$$v(t) \leqq 0, \qquad 0 \leqq w(t). \tag{4.8}$$

Denote by s_{left} and s_{right} the speed with which the shocks at the endpoints propagate; according to the jump relation (3.10).

$$s_{\text{left}} = \frac{f(v)-f(0)}{v}, \quad s_{\text{right}} = \frac{f(w)-f(0)}{w}. \tag{4.9}$$

Since a is an increasing function of $u, f(u)$ is convex. It follows from the mean value theorem that the difference quotient of f over an interval is not less than f' at the left endpoint, and not greater than f' at the right endpoint of that interval. So it follows from (4.8) that

$$a(v) \leqq \frac{f(v)-f(0)}{v}, \quad \frac{f(w)-f(0)}{w} \leqq a(w). \tag{4.10}$$

At this point we assume that a is strictly increasing, i.e., that for some positive number k

$$0 < k \leqq a'; \tag{4.11}$$

here we abbreviate d/du by prime. It follows that inequalities (4.10) are strict; combining these with (4.9) we can put them into this form

$$s_{\text{right}} - s_{\text{left}} \leqq \theta[a(w) - a(v)], \tag{4.12}$$

where θ *is* < 1.

Denote the length of $I(t)$ by $L(t)$; since s_{left} and s_{right} are the speeds with which the endpoints of I move,

$$\frac{dL}{dt} = s_{\text{right}} - s_{\text{left}}. \tag{4.13}$$

Substituting the inequalities (4.12) into (4.13) we get

$$\frac{dL}{dt} \leqq \theta[a(w) - a(v)].$$

Since by (4.8) $v < w$, $a(w) - a(v)$ is bounded by the total increasing variation $A^+(t)$ of $a(u)$ over $I(t)$:

$$a(w) - a(v) \leqq A^+(t). \tag{4.14}$$

Combining the last two inequalities we get

$$\frac{dL}{dt} \leqq \theta A^+(t).$$

Using inequality (4.7) we get

$$\frac{dL}{dt} \leqq \frac{\theta}{t} L(t);$$

and multiplying by $t^{-\theta}$ we deduce that

$$\frac{d}{dt}(t^{-\theta}L) \leqq 0.$$

Thus $t^{-\theta}L(t)$ is a decreasing function of time; in particular

$$L(t) \leqq t^{\theta}L(1) \qquad \text{for } t > 1. \tag{4.15}$$

Substituting this into the right side of (4.7) we get

$$A^+(t) \leqq t^{\theta-1}L(1).$$

Since $\theta < 1$, this shows that $A^+(t) \to 0$ as $t \to \infty$.

It follows from the strictly increasing character (4.11) of $a(u)$ that the total increasing variation of u along $I(t)$ is bounded by $A^+(t)/k$. Since u is $\leqq 0$ at the left endpoint of $I(t)$ and $\geqq 0$ at the right endpoint, it follows that likewise *the maximum* $m(t)$ *of* $u(x,t)$ *over* $I(t)$ *is bounded by* $A^+(t)/k$;

$$m(t) \leqq A^+(t)/k. \tag{4.16}$$

Combining this with the above estimate for A^+ we get that $m(t) \leqq \text{const}\; t^{\theta-1}$ which shows that *the maximum of* u *at time* t *tends to zero like* $t^{\theta-1}$.

This result is somewhat crude; a more detailed analysis will furnish a more precise result. (A different derivation was given by Barbara Quinn in her dissertation at New York University, 1970.) We start by expressing $f(r)$, $f(w)$ in (4.9) by their Taylor expansions; we get

$$\begin{aligned} s_{\text{left}} &= f'(0) + \frac{1}{2}f''(0)v + \frac{1}{6}f'''(\bar{v})v^2, \\ s_{\text{right}} &= f'(0) + \frac{1}{2}f''(0)w + \frac{1}{6}f'''(\bar{w})w^2, \end{aligned} \tag{4.17}$$

where $v < \bar{v} < 0,\ 0 < \bar{w} < w$.

Denote by K an upper bound for f'''; since m is an upper bound for $|v|$ and w, it follows that

$$s_{\text{left}} \geqq f'(0) + \frac{1}{2}\left[f''(0) + \frac{K}{3}m\right]v$$

$$s_{\text{right}} \leqq f'(0) + \frac{1}{2}\left[f''(0) + \frac{K}{3}m\right]w.$$

Substituting this into (4.13) we get

$$\frac{dL}{dt} \leqq \frac{1}{2}\left[f''(0) + \frac{K}{3}m\right](w - v). \tag{4.18}$$

It follows from (4.11) and (4.14) that

$$w - v \leqq \frac{a(w) - a(v)}{k} \leqq \frac{A^+(t)}{k}. \tag{4.19}$$

The constant k in (4.11) has to be a lower bound of $a' = f''(u)$ for $|u| \leqq m$; in particular we can take

$$k = f''(0) - Km. \tag{4.20}$$

Substituting this into (4.19) and then into (4.18) we get that for m small enough

$$\frac{dL}{dt} \leqq \frac{1}{2}\left[\frac{f''(0) + K/3\,m}{f''(0) - Km}\right]A^+ \leqq \left(\frac{1}{2} + Hm\right)A^+. \tag{4.21}$$

We substitute into (4.21) estimate (4.16) for m, and then estimate (4.7) for A^+; we obtain the following inequality:

$$\frac{dL}{dt} \leqq \left(\frac{1}{2} + \frac{H}{k}\frac{L}{t}\right)\frac{L}{t}. \tag{4.22}$$

Introduce a new variable J by $L = J\sqrt{t}$; (4.22) becomes

$$\sqrt{t}\,\frac{dJ}{dt} \leqq \frac{H}{k}\frac{J^2}{t}.$$

Dividing by $\sqrt{t}\,J^2$ we get, after integrating from T to $t > T$, that

$$\frac{1}{J(T)} - \frac{1}{J(t)} \leqq \frac{H}{2k}\left(\frac{1}{\sqrt{T}} - \frac{1}{\sqrt{t}}\right),$$

which implies that

$$\frac{1}{J(T)} - \frac{H}{2k\sqrt{T}} \leqq \frac{1}{J(t)}. \tag{4.23}$$

According to (4.15), $L(T)/T = J(T)/\sqrt{T}$ tends to 0 as $T \to \infty$; this implies that

for T large enough, the left side of (4.23) is positive. Then (4.23) furnishes an upper bound for $J(t)$ for all $t > T$. The boundedness of $J(t)$ implies that $L(t)$ is $O(\sqrt{t})$ as $t \to \infty$. Combining this with the estimates (4.7) and (4.16) we reach the following conclusion.

THEOREM 4.1. *Let u be a possibly discontinuous solution of the conservation law $u_t + f_x = 0$, where f is three times differentiable and strictly convex. Suppose that all discontinuities of u satisfy* (3.13), *and that $u(x,0)$ has compact support. Then*

(a) *the length of the support of $u(x,t)$ is $O(\sqrt{t})$,*

(b) $\mathrm{Max}_x \left| u(x,t) \right| = O(1/\sqrt{t})$.

It turns out that this result is rather precise: Using an explicit formula one can show, see [9], that the length of the support of u divided by $\sqrt{t}$ tends to a limit, and so does $\sqrt{t}\ \mathrm{Max}\ |u|$.

We turn now to solutions which are **periodic** in x:

$$u(x + p, t) = u(x,t).$$

We take $I(T)$ to be any interval of length p at time T. According to our basic estimate (4.7), the increasing variation of $a(u)$ per period is $\leqq p/T$. It follows then from (4.11) that the increasing variation per period of u itself does not exceed p/kT. Since u is periodic, its decreasing and increasing variations are equal, and serves as bound for the oscillation of u, in particular for the deviation of u from its mean value per period.

For a periodic solution $u(x,t)$, the flux f at $(0,t)$ equals the flux at (p,t); thus the total flux into an interval of length p is zero, and so the mean value of u,

$$\bar{u} = \frac{1}{p} \int_0^p u(x,t)dx,$$

is independent of t. We summarize our results as follows:

THEOREM 4.2. *Let $u(x,t)$ be a possibly discontinuous solution of $u_t + f_x = 0$, f strictly convex, $f'' > k > 0$. Suppose that all discontinuities of u satisfy* (3.13) *and that u is periodic in x with period p. Then*

(a) *The total variation of u at time t does not exceed $2p/kt$,*

(b)

$$\left| u(x,t) - \bar{u} \right| \leqq 1/kt, \tag{4.24}$$

where $\bar{u}$ is the mean value of u.

Again it can be shown that (4.24) is sharp, i.e., that

$$\lim_{t \to \infty} t \max_x \left| u(x,t) - \bar{u} \right| = k = f''(\bar{u}). \tag{4.25}$$

The surprising, almost paradoxical feature of inequality (4.24) is that it holds

uniformly for all solutions with period p; it is independent of the amplitude of the initial disturbance. All that the initial amplitude can influence is the time when the asymptotic estimate (4.24) becomes accurate: The *larger* the initial amplitude, the *sooner* (4.25) converges. This is in sharp contrast to the linear case where the asymptotic amplitude of a signal for large time is proportional to its initial amplitude, but the time it takes to reach the asymptotic shape is independent of the initial amplitude.

Let $u_1(x)$ be an initial function which is zero outside the interval $[0,p]$, and define $u_2(x)$ to be equal $u_1(x)$ in $[0,p]$, and periodic (see Fig. 8).

According to Theorem 4.1, $u_1(x,t)$ decays like $1/\sqrt{t}$; $u_2(x,t)$ on the other hand is periodic,[1] so its asymptotic behavior is governed by Theorem 4.2: $u_2(x,t)$ decays like $1/t$. So we have the paradoxical result that u_2, which represents a much larger initial disturbance than u_1, nevertheless decays faster than u_1.

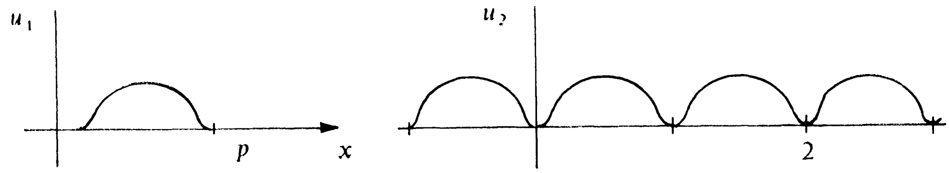


FIG. 8

5. Systems of conservation laws. Models which are at all realistic are governed by a whole system of conservation laws, rather than by a single one. The value of what we have learned about single equations lies in the light this knowledge sheds on systems. It turns out that the main phenomena we have found: the breakdown of continuous solutions, the necessity of imposing an entropy-like condition to distinguish those discontinuous solutions which are physically realizable from those which are not, and the decay of solutions as $t \to \infty$, have their counterparts for systems. That is not to say that the theory is as far advanced for systems as it is for single equations; on the contrary, what we have is a sea of conjectures, confined partly by the shores of numerical computations, with a few islands of solidly proved mathematical facts.

What are the proven facts about systems? In [**10**] the author has shown that solutions of 2×2 systems of conservation laws break down after a finite time, unless the initial data satisfy a monotonicity condition. In [**9**], an analogue of the entropy condition (3.13) is described, and a condition for genuine nonlinearity is given. In [**15**], Oleinik gives a uniqueness theorem for solutions of systems of two conservation laws of which one is linear. In [**2**], Glimm solves the initial value problem for systems, for initial data with small oscillation. In [**5**], Johnson and Smoller solve the initial value problem for initial data which satisfy a certain monotonicity condition, for 2×2 systems which satisfy a certain convexity-like condition. The only existence

[1] Solutions whose initial values are periodic are periodic for all t; this follows from the uniqueness theorem that solutions which are equal at $t = 0$ are equal for all $t > 0$.

theorem with no restrictions on the initial data is due to Nishida, [12], and works only for the system

$$u_t + v_x = 0, \qquad v_t - \left(\frac{1}{u}\right)_x = 0.$$

In [3], Glimm and the author prove the decay of solutions with small oscillation of 2×2 systems. The method described in Section 4 is taken from that paper.

For those who wish to work in this field I recommend Glimm's paper [2]. It contains a wealth of ideas, such as the use of an approximation scheme containing a sequence of random parameters; the scheme is shown to converge for almost all values of the parameters. Glimm also introduces novel, nonlocally defined functionals; the estimate of the growth and decay of these functionals plays a crucial role in the existence theorem.

This article is an expanded version of an invited address delivered at the January 1970 meeting of the MAA at San Antonio, Texas. Other versions of this talk were given at Oregon State University, Corvallis; Texas Tech. University, Lubbock, and at Brown University. The talk is partly based on the joint paper [3] with James Glimm.

References

1. A. Douglis, An ordering principle, Comm. Pure Appl. Math., 12 (1959) 87.

2. J. Glimm, Solutions in the large for nonlinear hyperbolic systems of equations, Comm. Pure Appl. Math., 18 (1965) 697–715.

3. J. Glimm, and P. D. Lax, Decay of solutions of systems of nonlinear hyperbolic conservation laws, Mem. Amer. Math. Soc., No. 101 (1970).

4. E. Hopf, The partial differential equation $u_t + uu_x = \mu u_{xx}$, Comm. Pure Appl. Math., 3 (1950) 201–230.

5. J. L. Johnson, and J. Smoller, Global solutions for an extended class of hyperbolic systems of conservation laws, Arch. Rational Mech. Anal., 32 (1969) 169–189.

6. S. Iv. Krushkov, Results on the character of continuity of solutions of parabolic equations and some of their applications, Mat. Zametki, 6 (1969) 97–108.

7. ———, First order quasi-linear equations in several independent variables, Math. USSR Sbornik, 10 (1970) No. 2.

8. P. D. Lax, Weak solutions of nonlinear hyperbolic equations and their numerical computation, Comm. Pure Appl. Math., 7 (1954) 159–193.

9. ———, Hyperbolic systems of conservation laws, II, Comm. Pure Appl. Math., 10 (1957) 537–566.

10. ———, Development of singularities of solutions of nonlinear hyperbolic partial differential equations, J. Mathematical Phys., 5 (1964) 611–613.

11. ———, On a notion of entropy, Proc. of Symposium at the University of Wisconsin, 1971, ed. E. Zarantonello.

12. T. Nishida, Global solutions for an initial value problem of a quasilinear hyperbolic system, Proc. Japan Acad., 44 (1968) 642–646.

13. O. A. Oleinik, Discontinuous solutions of nonlinear differential equations, Uspehi Mat. Nauk, (1957) 3–73, English Translation in Amer. Math Soc. Trans., Ser. 2, No. 26, pp. 95–172.

14. ———, On the uniqueness of the generalized solution of the Cauchy problem for a nonlinear system of equations occurring in mechanics, Uspehi Mat. Nauk, 78 (1957) 169–176.

15. ———, Uspehi Mat. Nauk (N. S.), 14 (1959) 165–170.

16. B. Quinn, Solutions with shocks, an example of an L_1-contractive semigroup, Comm. Pure Appl. Math., 24 (1971).

INFINITESIMALS

A. H. LIGHTSTONE, Queen's University

1. Introduction. The goal of this article is to enliven Abraham Robinson's concept of an infinitesimal by exhibiting infinitesimals in a simple and direct manner. Robinson has shown that the real number system $\mathscr{R}$ can be extended to a number system $\mathscr{R}^*$ that includes both infinitely large and infinitely small numbers. This he achieved by postulating the existence of a positive number that is less than each positive real number, and taking as additional postulates all statements that are true for $\mathscr{R}$. From a historical viewpoint establishing the existence of $\mathscr{R}^*$ is an enormous achievement. From a pragmatic viewpoint this piece of pure mathematics is outstanding for its capacity to revolutionize elementary mathematics by eliminating certain Weierstrass epsilon-delta statements that purport to define such basic concepts as *continuity* and *uniform continuity*. Instead, using infinitesimals, we can express the intuitive ideas involved in a direct and simple manner (contrast (1) and (2) below). For these reasons it is worthwhile to improve our appreciation of infinitesimals. To this end we shall exploit the familiar notion of the decimal expansion of a real number. After all, this notion improves our grasp of real numbers; just so, it will help us appreciate infinitesimals, indeed all numbers of our extended number system.

Now, the notion of a decimal expansion involves the natural number system, to be specific it involves a mapping of N into the digits $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$; as well we need the concept of an infinite sum. Accordingly, a decimal expansion in $\mathscr{R}^*$ involves the extended natural number system, in particular a mapping of N* into the digits. We shall need some facts about infinite natural numbers, so we begin by sketching the extended natural number system.

2. Extended natural number system. Using the *Compactness Theorem* of mathematical logic (which asserts that a set of statements has a model if each of its finite subsets has a model) it is easy to prove that the natural number system $\mathscr{N} = (\mathrm{N}, +, \cdot, <, 1)$ can be extended to a number system $\mathscr{N}^* = (\mathrm{N}^*, +, \cdot, <, 1)$ that possesses each algebraic property of $\mathscr{N}$. This is achieved by forming an enormously large postulate-set, contrary to the usual practice of minimizing the size of a postulate-set. The idea is to take as postulates all statements that are true for $\mathscr{N}$ and the statements $\omega > 1$, $\omega > 2$, $\omega > 3$, and so on; i.e., we postulate $\omega > n$ whenever

Professor Lightstone received his Toronto Ph. D. under Abraham Robinson. He has held positions at the University of Alberta, University of California, Berkeley, Carleton University, Victoria College, and Queen's University. He is currently on sabbatical leave at Yale on a Canadian Council Fellowship. His main research is foundations, and he is the author of *The axiomatic method; an introduction to mathematical logic* (Prentice-Hall 1964), *Concepts of calculus* I, II (Harper & Row 1965–66), *Symbolic logic and the real number system* (Harper & Row 1965), *Fundamentals of linear algebra* (Appleton-Century-Crofts 1969), and *Linear algebra* (Appleton-Century-Crofts 1969). *Editor.*

$n \in$ N. Clearly, each finite subset of this postulate-set has a model, e.g., $\mathcal{N}$ itself. So, by the *Compactness Theorem*, the postulate-set has a model, which we call $\mathcal{N}^*$. The term "statement" refers to wff of a Predicate Calculus built around $\mathcal{N}$; roughly, a statement involves only concepts of $\mathcal{N}$ and logical connectives, moreover each quantifier refers to N, the supporting set of $\mathcal{N}$. Actually, we can relax these requirements by the simple device of adjoining more terms to the tuple $\mathcal{N}$; these terms represent concepts of the natural number system, e.g., the set of all primes, the set T of all finite tuples whose terms are natural numbers, the operation S of summing the terms of a finite tuple. Moreover, we can allow quantification over several terms of $\mathcal{N}$, i.e., we regard $\mathcal{N}$ as possessing several supporting sets. The technical details are the concern of mathematical logic; here we accept the fact that $\mathcal{N}^*$ is an extension of an enriched natural number system $\mathcal{N}$, that $\mathcal{N}^*$ involves a number ω as postulated, and that each algebraic property of $\mathcal{N}$ (i.e. each statement true for $\mathcal{N}$) is true for $\mathcal{N}^*$ when interpreted in $\mathcal{N}^*$. Here $\mathcal{N}$ is the enriched natural number system built from $(\mathrm{N}, +, \cdot, <, 1)$ by incorporating terms that represent additional concepts of the natural numbers.

The key fact about N* is that this set is a superset of N and each of its members is finite or infinite. Now, an **infinite** natural number ∞ has the property that $\infty > n$ whenever $n \in$ N. So, a *finite* number t is less than some natural number. It can be shown that the finite natural numbers are the members of N, and the infinite natural numbers are the members of N* $-$ N. Our technique for proving statements of this sort relies on the fact that $\mathcal{N}^*$ possesses each algebraic property of $\mathcal{N}$; so to establish a fact about $\mathcal{N}^*$ we must find a fact about $\mathcal{N}$ which when interpreted in $\mathcal{N}^*$ yields the desired conclusion.

The point that we wish to emphasize is this. Whereas the natural numbers are given by the list $1, 2, 3, \cdots$ the extended natural numbers are given by the list

$$1, 2, 3, \cdots; \cdots, \infty, \infty + 1, \infty + 2, \cdots,$$

where ∞ is an infinite natural number. Notice our use of a semicolon as a **separator**, separating the finite natural numbers from the infinite natural numbers. We point out that the two parts of this list are *not* of the same sort. The part to the left of the semicolon, which exhibits the finite natural numbers, has the property that between any two of its terms there are only a finite number of numbers. The part to the right of the semicolon, which exhibits the infinite natural numbers, does not have this property; i.e., there are infinitely many numbers between two of its terms (if appropriately chosen), e.g., there are infinitely many numbers between ∞ and $\infty + \infty$.

Notice that there is no smallest infinite natural number; this follows from the fact that each natural number, except 1, has an immediate predecessor.

3. Extended real number system. The procedure that allows us to extend the natural number system to a number system that includes infinite numbers, also allows us to extend the real number system to a number system that includes both infinitely

large and infinitely small numbers. First, we mention that by the real number system $\mathscr{R}$ we mean the infinite tuple whose first term is R, the set of all real numbers, and whose remaining terms represent concepts of the real numbers (e.g., a specific mapping of R into R, the set of all finite tuples whose terms are real numbers, the set of all natural numbers). Next, we form our postulate-set. This consists of all statements (from our restricted language) that are true for $\mathscr{R}$ plus certain other statements that collectively assert the existence of a greater number than each natural number; namely the statements

$$\omega > 1, \omega > 2, \omega > 3, \cdots.$$

Clearly $\mathscr{R}$ is a model of each finite subset of this postulate-set; so, by the *Compactness Theorem*, this set of statements has a model which we call $\mathscr{R}^*$. We can choose $\mathscr{R}^*$ so that each of its first-order terms is a superset of the corresponding term of $\mathscr{R}$ (a *first-order* term of a number system is a set of numbers or tuples of numbers).

By virtue of its construction $\mathscr{R}^*$ is an extension of $\mathscr{R}$ that involves an infinite number ω; moreover, each statement that is true for $\mathscr{R}$, is true for $\mathscr{R}^*$ when interpreted in the language of $\mathscr{R}^*$. This means that each concept of the real number system extends to a corresponding concept of $\mathscr{R}^*$ that possesses all algebraic properties of the concept in $\mathscr{R}$. So N extends to N*; the set T of all finite tuples of real numbers extends to T* which contains tuples of infinite length; Σ which sums the terms of each tuple in T, extends to Σ^* which sums the terms of each tuple in T*.

Notice that the powerful proof-technique that so easily allows us to verify facts about $\mathscr{N}^*$, is available to verify facts about $\mathscr{R}^*$. To illustrate this, recall that each non-zero real number has a multiplicative inverse; this is true for $\mathscr{R}$ so it is true for $\mathscr{R}^*$ when interpreted in $\mathscr{R}^*$. This means that each nonzero member of R* has a multiplicative inverse. In particular, $\omega \neq 0$ so ω has a multiplicative inverse $1/\omega$. Now, by an *infinitesimal* we mean a member of R*, say ε, such that $|\varepsilon| < h$ whenever h is a positive real number (here, *absolute value* and *less than* are extensions to $\mathscr{R}^*$ of concepts of $\mathscr{R}$). It is easy to verify that $1/\omega$ is an infinitesimal. Indeed, the multiplicative inverse of each infinite number is an infinitesimal. By an *infinite* number we mean any number ∞ such that $|\infty| > h$ whenever h is real. Of course, by a *number* we mean a member of R*; we shall practice this convention hereafter.

Next, following Robinson, we introduce an important equivalence relation on R*. Let $a \in$ R* and $b \in$ R*; we say that $a \simeq b$ (read "a approximates b") if $a - b$ is an infinitesimal. Now, 0 is an infinitesimal, $-\varepsilon$ is an infinitesimal if ε is an infinitesimal, and the sum of two infinitesimals is an infinitesimal; so our relation $\simeq$ is an equivalence relation on R*. This equivalence relation allows us to express the idea that a number is *approximated* by another number (or that a point is *close* to another point). Of course, this is what calculus is all about. For example, when we say that a function f is continuous at a real number a, we mean:

(1) $f(x)$ approximates $f(a)$ whenever x approximates a, which is a statement about

$\mathscr{R}^*$. There is no way that we can express the idea that one real number approximates another real number within the real number system; this is a shortcoming of $\mathscr{R}$. Hence, there is no way that we can express (1) within the real number system. Instead, we use a Weierstrass epsilon-delta statement such as:

(2) Corresponding to each positive real number ε, there is a positive real number δ such that $|f(x) - f(a)| < \varepsilon$ whenever $|x - a| < \delta$.

Of course, (2) does not express the idea contained in (1); instead, it is merely equivalent to (1) in the sense that both statements are true or both are false. More precisely, (2) is true for $\mathscr{R}$ if and only if (1) is true for $\mathscr{R}^*$.

So far we have defined what is meant by an *infinite* number, what is meant by an *infinitesimal,* and we have defined the equivalence relation $\simeq$. Next, we present a basic fact about $\mathscr{R}^*$ that involves finite numbers; a *finite* number is a member of R* that is not infinite, so a is finite if there is a real number h such that $|a| < h$. The following fact shows that there is a stronger connection between R and the finite numbers than we might suspect at first sight.

FUNDAMENTAL THEOREM ABOUT FINITE NUMBERS. *Each finite number is approximated by a unique real number.*

Proof: Let t be any finite number and let $K = \{y \mid y \in \mathrm{R} \text{ and } y < t\}$. By the *Completeness Theorem* for $\mathscr{R}$, K has a least upper bound, say a. We claim that a approximates t. If not, there is a positive real number h such that $h < |t - a|$. There are just two cases since $a \neq t$:

(1) Assume $a < t$. Then $h < t - a$, so $h + a < t$, thus $h + a \in K$. But $a < a + h$; thus a is not an upper bound of K.

(2) Assume $t < a$. Then $h < a - t$, so $t < a - h$. Thus $a - h$ is an upper bound of K; but $a > a - h$, so a is not the least upper bound of K.

This proves that $a \simeq t$. Of course, if a finite number is approximated by two real numbers, then the difference of these real numbers is an infinitesimal, which can only be zero; so the real numbers are the same.

Incidentally, this proves that each interval of infinitesimal length contains at most one real number. For example, the open interval $(a - \varepsilon, a + \varepsilon)$ where $a \in \mathrm{R}$ and ε is a positive infinitesimal, contains exactly one real number; the open interval $(\varepsilon, 2\varepsilon)$ contains no real number.

4. Decimal expansions. By construction, $\mathscr{R}^*$ possesses both infinitely large and infinitely small numbers. Our goal, however, is to exhibit infinitesimals in a direct and convincing manner, not merely to prove that they exist. The idea is simple. Each real number has a unique decimal expansion, so each member of R* has a unique decimal expansion. In particular, if $0 < x < 1$, $x \in \mathrm{R}$, then x has the form $.d_1d_2d_3\cdots$, where each d_i is a digit. This means that there is a mapping d of N into the digits such that $x = \Sigma_{\mathrm{N}} d_n/10^n$, where $d_n = d(n)$. Now, this is true for $\mathscr{R}^*$ when interpreted in $\mathscr{R}^*$. So, each member of R* between 0 and 1, say y, has a

decimal expansion, i.e. there is a mapping d of N^* into the digits such that $y = \sum_{N^*} d_n/10^n$, i.e. $y = .d_1 d_2 d_3 \cdots ; \cdots d_\infty \cdots$. Notice our use of a semicolon to separate the terms of our sum that correspond to finite natural numbers, from terms that correspond to infinite natural numbers.

Now, let ∞ be a specific infinite natural number; certainly 10^∞ has a multiplicative inverse whose decimal expansion is

$$.000\cdots ; \cdots 000 \cdots 0\hat{1}0 \cdots$$

where the $\wedge$ indicates the ∞-th place.

Here we are involved with the mapping that associates 1 with ∞ and associates 0 with all other members of N^*. We have succeeded in our goal of exhibiting an infinitesimal.

We mention that each member of R^* whose decimal expansion has the form $.000\cdots ; \cdots d_\infty \cdots$ is an infinitesimal (here each digit to the left of the semicolon is 0). A word of caution. Whereas $\sum_N d_n/10^n \in R$ whenever d is a mapping of N into the digits, it is not true that $\sum_{N^*} d_n/10^n \in R^*$ whenever d is a mapping of N^* into the digits. For example, the decimal expansion $.000 \cdots ; \cdots 999 \cdots$ is *not* a member of R^*. If it is, then call it x. Clearly x is an infinitesimal since $|x| < 1/10^n$ whenever $n \in N$, so $|x| < h$ whenever h is a positive real number. Let ∞ be any infinite natural number, so $1/10^\infty$ is an infinitesimal, indeed a positive infinitesimal. Now, the sum of two infinitesimals is an infinitesimal, so $x + 1/10^\infty$ is an infinitesimal. Moreover, $x + 1/10^\infty > x$ so the decimal expansion of $x + 1/10^\infty$ has the form $.d_1 d_2 d_3 \cdots ;$ $\cdots d_\infty \cdots$ where a digit to the left of the semicolon is not zero. This means that $x + 1/10^\infty$ is not an infinitesimal. In view of this contradiction, we conclude that $.000\cdots ; \cdots 999 \cdots \notin R^*$.

To resolve this paradox, let M be the set of all mappings of N into $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Now M^*, the interpretation of M in $\mathscr{R}^*$, is a certain set of mappings of N^* into $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Our paradox merely proves that M^* does not contain all mappings of N^* into the digits. Thus, the statement "each decimal expansion represents a number" is true for $\mathscr{R}^*$ provided we interpret it correctly, that is, we must restrict the notion of a decimal expansion to certain, but not all, mappings of N^* into the digits.

We have already mentioned that a number is an infinitesimal provided its decimal expansion involves only zeros to the left of the semicolon. In line with this observation it is plausible to conjecture that if $x = .d_1 d_2 d_3 \cdots ; \cdots d_\infty \cdots$ is a number, then

$$x = .d_1 d_2 d_3 \cdots ; \cdots 0 \cdots + .000 \cdots ; \cdots d_\infty \cdots,$$

where $.d_1 d_2 d_3 \cdots ; \cdots 0 \cdots$ is real and $.000\cdots ; \cdots d_\infty \cdots$ is an infinitesimal. Remember that the *Fundamental Theorem about Finite Numbers* assures us that there is a unique real number a and a unique infinitesimal ε such that $x = a + \varepsilon$. Here, we conjecture that $a = .d_1 d_2 d_3 \cdots ; \cdots 0 \cdots$ and $\varepsilon = .000 \cdots ; \cdots d_\infty \cdots$. To see that this

conjecture is false, consider the multiplicative inverse of 3. Since $1/3 = .333\cdots$ in $\mathscr{R}$, it follows that $1/3 = .333\cdots;\cdots 3\cdots$ in $\mathscr{R}^*$. So $1/3 = a + \varepsilon$ where $a = 1/3$ and $\varepsilon = 0$. Returning to our conjecture, if $.000\cdots;\cdots 3\cdots \in \mathrm{R}^*$ so does $3 \times .000\cdots;\cdots 3\cdots$ i.e., $.000\cdots;\cdots 9\cdots$ is a number. But we have already pointed out $.000\cdots;\cdots 9\cdots \notin \mathrm{R}^*$. Moreover, since the difference of two numbers is also a number, we conclude that neither $.000\cdots;\cdots 3\cdots$, nor $.333\cdots;\cdots 0\cdots$, is a number. Therefore we cannot break up the decimal expansion of 1/3 in the simple manner suggested by intuition.

5. More paradoxes. Our insight into $\mathscr{N}^*$ and $\mathscr{R}^*$ benefits by analysing more fallacious arguments. For example, we can regard Peano's *Induction Postulate* as an algebraic property of the natural number system if we incorporate $\mathscr{P}\mathrm{N}$, the set of all subsets of N, into $\mathscr{N}$ and allow quantification over $\mathscr{P}\mathrm{N}$. This means that the basic set of our structure is $\mathrm{N} \cup \mathscr{P}\mathrm{N}$ so that quantification is over $\mathrm{N} \cup \mathscr{P}\mathrm{N}$, and that each quantifier is relativized to either N or $\mathscr{P}\mathrm{N}$. For example, the *Induction Postulate* is

$$(3)\qquad \forall y[y \in \mathscr{P}\mathrm{N} \to (1 \in y \wedge \forall x[x \in \mathrm{N} \to (x \in y \to x' \in y)] \to y = \mathrm{N})].$$

The usual practice is that the scope of each quantifier is indicated typographically, so capital letters indicate quantification over $\mathscr{P}\mathrm{N}$ and lower case letters indicate quantification over N. Thus (3) is abbreviated by

$$(4)\qquad \forall S[1 \in S \wedge \forall x[x \in S \to x' \in S] \to S = \mathrm{N}].$$

Since (4) is true for $\mathscr{N}$, it is true for $\mathscr{N}^*$ when interpreted in $\mathscr{N}^*$. But N is a subset of N^* that meets the requirements of (4); so $\mathrm{N} = \mathrm{N}^*$, and it follows that $\mathscr{N} = \mathscr{N}^*$ (recall that N is a set of numbers and that $\mathscr{N}$ is the corresponding number system).

The fallacy in this argument is easy to spot. Certainly (4) is true for $\mathscr{N}^*$ when interpreted in $\mathscr{N}^*$. But (4) is an abbreviation for (3); so (3) is true for $\mathscr{N}^*$ when interpreted in $\mathscr{N}^*$. Each concept that appears in (3) must be interpreted in $\mathscr{N}^*$; in particular, $\mathscr{P}\mathrm{N}$ is interpreted in $\mathscr{N}^*$, i.e., there is a term of $\mathscr{N}^*$ called $(\mathscr{P}\mathrm{N})^*$. So

$$(5)\qquad \forall y[y \in (\mathscr{P}\mathrm{N})^* \to (1 \in y \wedge \forall x[x \in \mathrm{N}^* \to (x \in y \to x' \in y)] \to y = \mathrm{N}^*)]$$

is true for $\mathscr{N}^*$. In particular,

$$(6)\qquad \mathrm{N} \in (\mathscr{P}\mathrm{N})^* \to (1 \in \mathrm{N} \wedge \forall x[x \in \mathrm{N}^* \to (x \in \mathrm{N} \to x' \in \mathrm{N})] \to \mathrm{N} = \mathrm{N}^*)$$

is true for $\mathscr{N}^*$. The fallacy rests on the assumption that $\mathrm{N} \in (\mathscr{P}\mathrm{N})^*$, from which (6) allows us to conclude that $\mathrm{N} = \mathrm{N}^*$. In fact $\mathrm{N} \notin (\mathscr{P}\mathrm{N})^*$; indeed, $(\mathscr{P}\mathrm{N})^*$ is a certain collection of subsets of N^*, but not all, i.e., $(\mathscr{P}\mathrm{N})^* \neq \mathscr{P}(\mathrm{N}^*)$. The paradox proves this statement.

The same sort of unconscious slip accounts for our next paradox which revolves around the *Completeness Theorem* of the real number system, namely "Each non-empty set of real numbers which has an upper bound, also has a least upper bound."

To include this statement in our language we incorporate the following concepts as terms of $\mathscr{R} - \mathscr{P}\mathrm{R}$, upper bound of a set of real numbers, least upper bound of a set of real numbers. We take $\mathrm{US}a$ as an abbreviation for "a is an upper bound of S," and we take $\mathrm{LS}a$ as an abbreviation for "a is the least upper bound of S." In this language the *Completeness Theorem* is

$$\forall S[S \neq \varnothing \wedge \exists x \mathrm{US}x \rightarrow \exists y \mathrm{LS}y] \tag{7}$$

which expands to

$$\forall z[z \in \mathscr{P}\mathrm{R} \rightarrow (z \neq \varnothing \wedge \exists x(x \in \mathrm{R} \wedge \mathrm{U}zx) \rightarrow \exists y(y \in \mathrm{R} \wedge \mathrm{L}zy))]\,. \tag{8}$$

The paradox consists in observing that a structure that satisfies the *Completeness Theorem* does not possess infinitesimals, and hence a structure that possesses infinitesimals does not satisfy the *Completeness Theorem*. To prove this, let $\mathrm{S} = \{\varepsilon | \varepsilon \simeq 0\}$. Clearly S is nonempty and has an upper bound. Then, by the *Completeness Theorem*, S has a least upper bound, say t. But each member of R^* is either infinite or has the form $a + \varepsilon$, where $a \in \mathrm{R}$ and $\varepsilon \simeq 0$. In particular, t is not infinite, so $t = a + \varepsilon$. It is easy to see that $a = 0$ (otherwise $a/2 + \varepsilon$ is an upper bound of S). This means that an infinitesimal is the least upper bound of S; but this is also out of the question since $2t$ is an infinitesimal if t is an infinitesimal, and clearly $t < 2t$, so t is *not* an upper bound of S. We conclude that S does not have a least upper bound.

Returning to our fallacy notice that it is based on taking (8) at face value, i.e. failing to interpret (8) in $\mathscr{R}^*$. It is not the *Completeness Theorem* that is true for $\mathscr{R}^*$, rather it is the interpretation of the *Completeness Theorem* in $\mathscr{R}^*$ that is true for $\mathscr{R}^*$. Thus, from (8)

$$\forall z[z \in (\mathscr{P}\mathrm{R})^* \rightarrow (z \neq \varnothing \wedge \exists x(x \in \mathrm{R}^* \wedge \mathrm{U}^*zx) \rightarrow \exists y(y \in \mathrm{R}^* \wedge \mathrm{L}^*zy))] \tag{9}$$

is true for $\mathscr{R}^*$.

Again we must resist the assumption that $(\mathscr{P}\mathrm{R})^* = \mathscr{P}(\mathrm{R}^*)$; in fact, $(\mathscr{P}\mathrm{R})^*$ consists of certain subsets of R^*. However, $\{\varepsilon | \varepsilon \simeq 0\} \notin (\mathscr{P}\mathrm{R})^*$; indeed, the force of this paradox is to prove this fact.

6. The language of $\mathscr{R}$ and $\mathscr{R}^*$. The paradoxes show that the meaning of a statement about $\mathscr{R}$ can change in a subtle manner when the statement is interpreted in $\mathscr{R}^*$. Moreover, the fact that the postulate-set for $\mathscr{R}^*$ contains all statements, from a certain language, that are true for $\mathscr{R}$, provides us with a simple, yet powerful, method of proving facts about $\mathscr{R}^*$ by merely quoting appropriate and true statements about $\mathscr{R}$ (from the language involved). So, we need to understand the language that plays such an important role in the development of $\mathscr{R}^*$; we must pin down just what is meant by a "statement" in the context of the real number system.

The real number system involves a certain set of numbers, namely R the set of all real numbers; moreover, this number system involves many ideas or concepts

that can be represented mathematically by sets — sets whose members are real numbers, sets whose members are tuples of real numbers, sets whose members are sets of real numbers, etc. For example, the concept of a **natural number** is exemplified by the set of all natural numbers; the **less than** relation is exemplified by a certain set of pairs whose terms are real numbers; the binary operation of **addition** is exemplified by a certain set of triples whose terms are real numbers (more accurately, a set of pairs whose first terms are pairs); the operation of **summing** a list is exemplified by a set of tuples whose terms are real numbers (the last term represents the sum of the remaining terms of each tuple); the concept of a **finite tuple** is exemplified by a set whose members are tuples, where the terms of each tuple are real numbers; the concept of the **length** of a finite tuple is represented by a set whose members are pairs, the first term of each pair is a tuple, and each second term is a natural number; the notion of an **upper bound** of a set of real numbers is exemplified by a set whose members are pairs, first terms being sets and second terms being real numbers; the notion of a **set** of real numbers is characterized by the set of all subsets of R.

The first step in building up the language of $\mathscr{R}$ is to assign a name, i.e. a symbol, to each of its concepts. So "N" denotes the set of all natural numbers, "$<$" is a name for the **less than** relation on R, "$+$" is a name for the binary operation of addition on R, "S" is a name for the operation of summing a list that consists of real numbers; "T" denotes the set of all finite tuples whose terms are real numbers; "L" represents the concept of the length of a finite tuple of real numbers; "U" represents the concept of an upper bound of a set of real numbers; "$\mathscr{P}$R" represents the idea of a set of real numbers. Thus each of N, $<$, $+$, S, T, L, U, and $\mathscr{P}$R is a set. Of course, there are other concepts of $\mathscr{R}$ that will concern us; here, we have merely illustrated the idea that a concept of a number system is exemplified by a set and is denoted by a symbol.

Perhaps the most fundamental statement we can make concerning a set is that an object is a member of the set. We regard numbers, tuples, and sets as objects; these, together with the concepts of $\mathscr{R}$, generate statements such as $3 \in \mathrm{N}$, $(2, 5) \in <$, $(5, -1, 4) \in +$ which are true. On the other hand, the statements $-5 \in \mathrm{N}$, $(5, 2) \in <$, $(5, -1, 1) \in +$ are false. In the preceding sections of this article we have followed the usual custom of abbreviating a statement such as "$(a, b) \in <$" by writing "$a < b$", and of abbreviating "$(a, b, c) \in +$" by writing "$a + b = c$". A statement of the form $x \in \mathrm{S}$, where S is a concept of $\mathscr{R}$, is true provided that x indeed is a member of the set S that exemplifies the concept involved. Notice that our procedure for outlining the statements of our language (which we have not yet completed) also yields the truth-value of each statement; moreover, the problem of determining the truth-value of a certain statement can be reduced to the problem of deciding whether an object is a member of a certain set.

We build on the primitive, atomic statements obtained directly from the concepts of $\mathscr{R}$, by utilizing the connectives of symbolic logic. So, let p and q be any statements

of our language, either primitive statements just introduced above, or more complicated statements that are built up from atomic statements by means of our connectives. Then we say that "$\sim p$" (not p), "$p \vee q$" (p or q), "$p \wedge q$" (p and q) "$p \rightarrow q$" (if p then q), and "$p \leftrightarrow q$" (p if and only if q) are statements. Moreover, "$\sim p$" is true if p is false; "$p \vee q$" is false just in case both p and q are false; "$p \wedge q$" is true if both p and q are true; "$p \rightarrow q$" is false just in case p is true and q is false; "$p \leftrightarrow q$" is true if p and q have the same truth-value.

Next, let $P(x)$ be a statement-form, i.e. an expression involving a place-holder, here x, such that replacing x by an object yields a statement. Then we say that "$\forall x[P(x)]$" is a statement, moreover this statement is true provided that each statement generated by $P(x)$ is true. Similarly, "$\exists x[P(x)]$" is a statement; this statement is true provided that at least one of the statements generated by $P(x)$ is true. Each **quantifier** $\forall$ or $\exists$ must carry with it a set of objects used to generate statements from the statement-form involved. It is convenient, as an abbreviating device, to indicate the set of objects typographically, i.e. by using a special symbol for the place-holder that follows the quantifier. For example, lower case letters at the end of the alphabet (e.g., x, y, z) are used to indicate quantification over R; lower case letters at the middle of the alphabet (e.g., m and n) are used to indicate quantification over N; upper case letters (e.g., S and T) indicate quantification over $\mathscr{P}$R ; greek letters (e.g., α and β) indicate quantification over T , the set of all finite tuples.

We mention that each statement must be of finite length, i.e., each statement may contain only a finite number of instances of connectives (so, only a finite number of instances of atomic statements).

The language that we have just sketched is sufficiently rich and flexible to express the usual kind of statements that interest mathematicians. For example, the *Completeness Theorem* and the *Principle of Mathematical Induction*, as well as the postulates for an ordered field, fall within this language. Of course, if we have something to say about $\mathscr{R}$ we have only to search out the basic concepts involved and incorporate them as terms of the real number system. For this reason, we think of $\mathscr{R}$ as a number system involving many concepts (i.e. terms) some of which are specified, but not all.

An understanding of the extended number system $\mathscr{R}^*$ and its language requires a more sophisticated approach. Although the structure of this language conforms to the pattern for the language of $\mathscr{R}$ in the matter of how statements are constructed from given statements, it differs from the language of $\mathscr{R}$ when it comes to formulating the atomic statements of the language, the statements from which all statements of the language are constructed. The point is that we are not free to choose the concepts of $\mathscr{R}^*$, nor are we free to define their members arbitrarily. Instead, each concept of $\mathscr{R}^*$ is rooted in ac orresponding concept of $\mathscr{R}$, and is an extension of that concept (for first-order concepts this means it is a superset of the set that exemplifies the concept in $\mathscr{R}$). When we build a concept of $\mathscr{R}$ we can define it as we wish, i.e., we can choose

its members freely. However, when we *name* a particular term of $\mathscr{R}$ we must pay attention to the actual concept it represents. For example, if (1, 1, 3) is a member of a certain set, we do not call that set $+$.

Bear in mind that $\mathscr{R}^*$ is obtained axiomatically by forming the set of all statements that are true for $\mathscr{R}$, together with a set of statements that collectively postulate the existence of an infinite number. It is a giant step from a postulate-set to a *model* of that postulate-set (i.e. to prove its existence); of course, we cannot go into this here. However, we can point out that a model of this postulate-set is a number system patterned on $\mathscr{R}$ and involving an infinite number ω as postulated; i.e., each concept (term) of the postulated number system, except ω, corresponds to a concept of $\mathscr{R}$ and can be regarded as a set. Moreover, the language of $\mathscr{R}^*$, and the question of the truth-value of each statement of this language, is decided by the sets that represent the concepts of this number system in the same manner as for $\mathscr{R}$. This is where the distinction based on the *interpretation* in $\mathscr{R}^*$ of a concept of $\mathscr{R}$ enters the picture. The truth-value of a statement depends ultimately upon the sets that exemplify the concepts appearing in that statement. For most concepts, the set exemplifying a concept in $\mathscr{R}^*$ is not the set that represents it in $\mathscr{R}$. Moreover, and this is our main point, the set that represents a concept in $\mathscr{R}^*$ cannot be characterized verbally in the same direct and simple fashion as for $\mathscr{R}$. For example, the set of all subsets of R is *not* represented in $\mathscr{R}^*$ by the set of all subsets of R*; rather, it is exemplified by a certain set of subsets of R* (we have already proved this by way of a so-called paradox).

The idea of interpreting a statement in a number system is illustrated more simply by considering the statement $\forall x[x \neq 0 \rightarrow \exists y(xy = 1)]$. We are accustomed to interpreting this statement in several number systems, e.g., the real number system, the rational number system, and the system of integers. To determine its truth-value in a particular number system, we consider the operation of multiplication of that number system and the number set involved. We arrive at the conclusion that this statement is true for the real number system and for the rational number system, but is false for the system of integers. We have interpreted a statement in a number system, by interpreting the concepts involved in the statement, and have reached a decision regarding its truth-value in that number system.

References

1. Abraham Robinson, Topics in Non-Archimedean Mathematics, Symposium on the Theory of Models; North-Holland, Amsterdam, 1965.

2. Abraham Robinson, Non-Standard Analysis, North-Holland, Amsterdam 1966.

3. Abraham Robinson, Introduction to Model Theory and to the Metamathematics of Algebra, North-Holland, Amsterdam, 1963.

FIDELITY IN MATHEMATICAL DISCOURSE: IS ONE AND ONE REALLY TWO?

P. J. DAVIS, Brown University

"I wanted certainty in the kind of way in which people want religious faith. I thought that certainty is more likely to be found in mathematics than elsewhere. But I discovered that many mathematical demonstrations, which my teachers expected me to accept, were full of fallacies, and that, if certainty were indeed discoverable in mathematics, it would be in a new field of mathematics, with more solid foundations than those that had hitherto been thought secure. But as the work proceeded, I was continually reminded of the fable about the elephant and the tortoise. Having constructed an elephant upon which the mathematical world could rest, I found the elephant tottering, and proceeded to construct a tortoise to keep the elephant from falling. But the tortoise was no more secure than the elephant, and after some twenty years of very arduous toil, I came to the conclusion that there was nothing more that I could do in the way of making mathematical knowledge indubitable."

BERTRAND RUSSELL,
Portraits from Memory

1. Platonic mathematics. The twentieth century has not yet delineated definitively the working principles and the broad articles of faith of what has come to be called "Platonic mathematics". Among these principles might be listed:

1. *The belief in the existence of certain ideal mathematical entities such as the real number system.*
2. *The belief in certain modes of deduction.*
3. *The belief that if a mathematical statement makes sense, then it can be proven true or false.*
4. *The belief that fundamentally, mathematics exists apart from the human beings that do mathematics. Pi is in the sky.*

These beliefs have been questioned; and in the last century a number of distinguished mathematicians have raised their voices against one or more of them. These mathematicians include Kronecker, Borel, Brouwer, Gödel, Weyl, and in more recent times, E. Bishopp. One objection raised by some materialists is that the physical world may be completely finite, and this is hard to accommodate to an infinity of integers. Other objections have to do with the axiom of choice, the axiom of the excluded middle, etc.

Philip J. Davis received his Harvard Ph. D. under Ralph Boas. He has taught at Harvard, MIT, American University, Maryland, and Brown. He has extensive industrial and government experience, including five years as Chief, Numerical Analysis Section, National Bureau of Standards; also he was a Guggenheim Fellow in 1956–57. His extensive work in numerical analysis and applied mathematics includes the books *Lore of Large Numbers* (1961), *Interpolation and Approximation* (1963), *Mathematics of Matrics* (1964), *Approximate Numerical Integration* (with P. Rabinowitz, 1967), and 3.1416 *and All That* (with W. Chinn, 1969). Professor Davis received the 1960 Award in Mathematics of the Washington Academy of Sciences and the MAA Chauvenet Prize in 1963. *Editor.*

As far as No. 3 is concerned, the work of Gödel and the Logical School has put the *coup de grâce* on this principle; yet—and by no means strangely—it persists as a psychological prop in one's daily work. I once asked a very distinguished number theoretician whether he thought that Fermat's Last Theorem was one of the unprovable statements in the sense of Gödel. His answer was quick and definite: "It is not. We are just too dumb to find the proof." The truth of the matter is that if mathematics were ever to enter into a region where it is frustrated by too many interesting but unprovable statements, then this would cast a blight on the methodology and ritual surrounding the notion of proof.

The questioning of Platonic mathematics has led to other types of mathematics variously called intuitionistic mathematics, constructivistic mathematics, recursive mathematics, and other names. Some of these are subsets of the usual mathematics. The computing machine has undoubtedly reopened and reinforced some of the arguments. The reception given to non-Platonic mathematics ranges all the way from coolness to indifference. One recalls the story of Kronecker in the 1880's. Someone came to him and told him that Lindemann had just proved that pi was a transcendental number. "Very interesting," said Kronecker, "but pi doesn't exist." This skepticism was largely ignored. At a series of recent lectures on non-Platonic mathematics, a typical comment was "Well presented, but irrelevant. Let's get back to our (Platonic) drawing boards." Undoubtedly in 1971, one can earn a living with Platonic mathematics, and if mathematician A spouts some Platonism to mathematician B and the latter responds in kind, then there is at least human significance in the act. The emperor may be walking around in his underwear, but if the court is also, they can make a life together.

It is the object of this essay to present additional aspects of the non-Platonicity of mathematics.

Several years ago I did some experiments using the computer to prove and derive theorems in elementary analytic geometry, [2]. These experiments inevitably led to speculation on the difference in the level of credibility of a theorem which has been proved or derived by machine as opposed to one which has been "hand crafted" in the traditional fashion. This essay is an outcome of this experience. The particular arguments made here have not been put forth elsewhere at any length, and lead to the conclusion that mathematics, in some of its aspects, takes on the nature of an experimental science.

2. Symbols. It is commonplace that mathematics is done with symbols. Figures, words, graphs, special symbols of all sorts litter the mathematical page. The most common mode of operation is from the sheet of paper, the blackboard, the sandpit in the case of Archimedes, the TV computer screen in the case of a latter day Archimedes, into the brain through the eye and the optic nerve. Presumably, when this symbolic information enters the brain, it leaves a physical trace there. The symbols are then processed by the brain and hard copy output may be made via hand or

mouth. If there were never any oral or written or action output (such as with the educated horse who when cued stamps with his foreleg in answer to arithmetic problems) then mathematics might exist, but not in the manner in which we know it.

The principal symbol of mathematics, then, is the graphical symbol, perceived by the eye. There are blind mathematicians of first rank (such as L. Pontryagin) and it would be interesting to hear what he has to say about his manner of symbol formulation, manipulation, and space perception. I am not aware of any mathematicians who are blind and deaf mutes, but I presume that Helen Keller who graduated from Radcliffe could do sums.

If one believes in Platonic mathematics, then it is possible to free mathematics from the symbols that carry it. After all, the spoken word "two" and the Arabic symbol "2", the Braille symbol for two, have a common interpretation. Hence, there must be, so the argument goes, a concept of twoness which is symbol-free. As Plato put it, mathematical objects are perceived by the soul. Be this as it may, I cannot give a simple instance of symbolless, soul mathematics. Even if I knew one, how could I communicate it, short of telepathy?

3. Proof. One of our most precious inheritances from Greek mathematics is the notion of proof. Certain statements are derivable from other statements by means of "pure reason", and a corpus of connected material can be built up in which all statements are derived from a few fundamental statements known as axioms. This is the program set forth in Euclid, and this, after 2300 years, remains the beau ideal of mathematical exposition. In fact, some authorities believe that this is the hallmark of mathematics. Now, what is the purpose of a proof and how is a proof carried out? If you read Plato (Meno, 87) you find Socrates going through a derivation with a slave boy. Using the famous Socratic method, he leads the boy by the nose, so to speak, to the result that in a 45°, 45°, 90° triangle, the area of the square on the hypothenuse has double the area of the square on the short side. This dialogue creates the impression first of all of the derivation of new knowledge *ex nihilo* (or ex very little), and secondly of establishing firmly on the basis of a few easily accepted premises a statement which is far less transparent. To prove is to establish beyond the question of doubt, and mathematics has been thought capable of just such a thing. History does not prove, sociology does not prove, physics does not prove, philosophy does not prove, religion (if we can forget the church's unrequited seven hundred year love affair with Aristotelianism) does not prove. Mathematics alone proves, and its proofs are held to be of universal and absolute validity, independent of position, temperature or pressure. You may be a Communist or a Whig or a lapsed Muggletonian, but if you are also a mathematician, you will recognize a correct proof when you see one.

These two aspects of Socrates' teaching: proof as a program of certification—let's not call it establishing truth—and proof as a program of discovery and of new mathematics formation are present in today's mathematics. The most charming

instance of success of the first part of Euclid's program is undoubtedly contained in John Aubrey's brief life of the philosopher Thomas Hobbes:

> He (Thomas Hobbes) was 40 years old before he looked on Geometry; which happened accidentally. Being in a Gentleman's Library, Euclid's *Elements* lay open, and 'twas the 47 El. libri I. He read the Proposition. By G..., sayd he (he would now and then sweare an emphatical Oath by way of emphasis) this is impossible! So he reads the Demonstration of it, which referred him back to a Proposition, which Proposition he read. That referred him back to another, which he also read. *Et sic deinceps* [and so on] that at last he was demonstratively convinced of that trueth. This made him in love with Geometry.

But the facts of the matter are somewhat different. If you think you could talk to your favorite bartender and lead him by the nose à la Socrates and have him arrive at the Stone-Weierstrass theorem, think again. The path would turn him off the way I am turned off by Spinoza's proofs in ethics. As Poincaré observed, the ability to follow a mathematical argument is spread unevenly through the populace. For the professional mathematician, proof may be less a matter of convincing oneself psychologically of the truth of a statement than of merely assigning the tags 'true' or 'false' to the statement. But a balance must be struck. For as N. Bourbaki has written,

> "Indeed, every mathematician knows that a proof has not been 'understood' if one has done nothing more than verify step by step the correctness of the deductions of which it is composed and has not tried to gain a clear insight into the ideas which have led to the construction of this particular chain of deductions in preference to every other one."

Secondly, mathematics can and has been done in a "proofless" atmosphere. The Egyptians and Babylonians had piled up a considerable body of mathematics before even the Greeks came along with their proofs. If one reads Ptolemy one sees how proofless material can exist side by side with the mathematics of proof. In today's world, the physicist and engineer often work in absence of proof, it being sufficient to work formally and symbolically and have the work backed by a physical intuition or by an experimental confirmation.

Despite these two mathematical worlds, which have for a long time existed side by side, mathematicians, and in particular mathematical logicians have over the past century systematized and made precise the notion of a proof. Without attempting the technicalities, the matter seems to come down to this. The axioms, i.e., the primitive statements or assumptions are representable as certain strings of atomic symbols. The theorems are representable as certain other strings of atomic symbols. Proving is the process of passing from an axiom string to a theorem string by a finite sequence of allowable elementary transformations. To verify that the next

man's putative theorem is, in fact, the theorem he claims it to be, is merely to verify that the sequence of string transformations are in order. The whole thing is in principle perfectly mechanizable and is work for a slave boy or our modern equivalent, the computer. From this point of view to verify an advanced statement is similar to establishing the arithmetic theorem $123 + 456 = 579$. We merely process the data. Proof is at once the glory of mathematics and its least human aspect.

A proof can be compared with a program. The axioms are analogous to the input. The theorem is analogous to the output while the proof is the program. To find a proof consists of finding a program. To verify a given proof we need only rerun the program.

4. Fidelity. I come now to the nub of my argument. Mathematics, as we have seen, proceeds through symbols and symbol manipulation. It therefore assumes that we can create distinct symbols, recognize strings of symbols, reproduce symbols, concatenate symbols. A symbol has a physical trace. It is a blob of ink or a vibration in the air, etc. If I mark down two 1's these 1's may be identical on the macroscopic level, but not at the microscopic. It is impossible to create identical symbols. Like snowflakes, they are all different. If they are "nearly" identical, they may be perceived variously. The eye may be dim, the ear heavy, the brain fatigued. The computer may slip a pulse, its voltages may drop, it may be communicated with over a noisy channel.

As part of the assumptions of Platonic mathematics we should therefore list:

FIG. 1

Are all the symbols above instances of the same symbol?
As of 1971, high fidelity recognition by machine of hand written characters has proved to be difficult.

0. *Distinct Symbols can be Created. Instances of a given symbol can be created. Symbols can be processed and reproduced and concatenated with absolute fidelity. Symbols can be recognized as distinct or identical as the case warrants.*

An orthodox Platonist might say the above is unnecessary insofar as mathematics exists without physical carriers. A non-Platonist, particularly one who has been exposed to communication theory, will say this is nonsense. We can do these things only with a certain probability of success. The probability may be very high indeed, but there may be occasional failure. What is the mathematics of failure? Without making too many distinctions, let us agree indifferently to call an act of recognizing, reproducing, or processing one symbol 'an operation.' Let the probability of carrying out an operation with perfect fidelity be p. The number p satisfies the inequality

$$0 < p < 1$$

and we shall think of p as being very close to 1. A realistic value of p depends upon

who or what is doing the symbol processing and under what circumstances. I know that in doing sums or in typing up an IBM card my personal probability may be around

$$p \approx 1 - 10^{-2}.$$

I have heard figures around

$$p \approx 1 - 10^{-9} \text{ to } p \approx 1 - 10^{-12}$$

quoted for computing machines. Now if the probability of success in one elementary operation is p, then, assuming independence, which may or may not be true, the probability of success in a sequence of n operations is p^n. Thus if n is very large, this probability goes down considerably. Now what probability of failure will you tolerate? One in a thousand? Then you want

$$p^n \geqq 1 - 10^{-3} \quad \text{or} \quad n \log p \geqq \log (1 - 10^{-3}).$$

If now

$$p = 1 - \frac{1}{m},$$

then we want

$$n \leqq \frac{\log\left(1 - \frac{1}{1000}\right)}{\log\left(1 - \frac{1}{m}\right)}.$$

Since $\log(1 - h) \approx -h$ for small h, we need

$$n \leqq \frac{m}{1000}.$$

In other words, to keep within the required confidence limits, we should not carry out more than $m/1000$ operations. Now the number of operations which go on inside a computer are enormous, so that the chance of failure is not infinitesimal in terms of lifetime probabilities. (In "Computer Programming for Accuracy," Proceeding of the 1968 Army Numerical Analysis Conference, U. S. Army Research Office, Durham, North Carolina, J. M. Yohe lists 38 types of errors that may occur in carrying out a computer computation. These are grouped under seven major categories as follows: Errors due to hardware limitations, errors due to software limitations, errors due to hardware failure, errors due to software failure, errors due to program failure, errors due to faulty operation, errors due to inadequate planning. A similar list for mathematics produced in the conventional handcrafted fashion would surely be interesting.)

Repeating a computation by way of check helps, of course. If a complicated

computation is carried out with a probability of success of $1 - 1/r$ ($r > 1$), and is performed independently v times, then the probability of at least one success in the v blocks of computation is $1 - (1/r)^v$. Thus, the level of confidence is raised.

Consider then simple addition of numbers carried out in the usual way. If there are too many digits in the numbers, then the probability of a computation being accurate (or of discovering which of a block of independently arrived at answers is the correct one) might be small. The reader need only insert his favorite probabilities for himself and for his machine in the above formulas. Perhaps we need to take a number of over a million digits or over a billion digits to make success unlikely. No matter. Platonic mathematics guarantees an unlimited number of integers and each integer has a decimal representation.

Ordinary arithmetic is one of the most elementary of the mathematical disciplines. Among the theorems of arithmetic are the various sums. Here is a theorem in arithmetic: $12345 + 54321 = 66666$. If this theorem does not excite you particularly, this is your value judgement and is extraneous to the mathematical structure. It might excite a Kabalist or an income tax consultant. Now, as we have observed, the arithmetic of excessively large numbers can be carried out only with diminishing fidelity. As we get away from trivial sums, arithmetic operations are enveloped in a smog of uncertainty. The sum $12345 + 54321$ is not 66666. It is not a number. It is a probability distribution of possible answers in which 66666 is the odds-on favorite. (A somewhat less transparent example is this. Consider the popular solitaire game called "Canfield". If the rules are fixed, and the line of play specified unambiguously, then the expected value of Canfield constitutes a mathematical theorem which is of considerable interest in some quarters. As far as I am aware, because of the complexity of Canfield, no one has been able to use the elementary textbook theorems on combinatorial probability to arrive at the expected value. Yet, all we have to do in principle is to examine each of the 52! games that are possible and average their values.)

There is a parallel with the limitations of physical measurement. There is wisdom in the primitive counting system one, two, three, many, myriads.

PROBLEM: Given

$A =$ 11777777711171717171777171171711111177717177711771177171717171777171777171717
171777111717111111717777111717171111717177171

$B =$ 7777717117111177777771111111117717171711177777171777711171711111717117171777
11111117171777777777111717177771111777117177771

Find $A + B$.

FIG. 2

The numbers A and B cannot be reproduced with perfect fidelity, let alone added.

5. Fidelity in proofs. The authenticity of a mathematical proof is established by verifying that a sequence of transformations of atomic symbol strings is legitimate. In point of fact, proofs are not written in terms of atomic strings. They are written in a mixture of common discourse and mathematical symbols. *Definitions* are made to serve as abbreviations for longer combinations of words and symbols. *Lemmas* are introduced as temporary platforms and scaffoldings from which one can argue with less fatigue and hence greater security. *Corollaries* are introduced for the psychological lift of obtaining deep theorems cheaply.

Splicing two theorems is standard practice. In the course of a proof, one cites Euler's Theorem, say, by way of authority. The onus is now on the reader to supply the particular theorem of Euler that the author is talking about and to verify that all the conditions (in their most modern formulation) which are necessary for the applicability of the theorem are, in fact, present.

If splicing is common to lend authority, then *skipping* is even more common. By skipping, I mean the failure to supply an important argument. Skipping occurs because it is necessary to keep down the length of a proof, because of boredom (you cannot really expect me to go through every single step, can you?), superiority (the fellows in my club all can follow me) or out of inadvertence. Thus, far from being an exercise in reason, a convincing certification of truth, or a device for enhancing the understanding, a proof in a textbook on advanced topics is often a stylized minuet which the author dances with his readers to achieve certain social ends. What begins as reason soon becomes aesthetics and winds up as anaesthetics.

To go from the foundations of mathematics to any of the advanced topics on the frontier can be done in about 5 or 6 books. Perhaps 1500 pages of proof text of current style. This is humanely broken into smaller bits. The lengths of these smaller bits vary from discipline to discipline. Perhaps number theory has the longest individual proofs. I know one proof in Landau which is over a hundred pages long. I have before me a book on advanced topics in analysis just off the press. The average length of the proofs seems to be about 10 lines. This mirrors the *sitzfleisch* of the contemporary reader.

I do not know many people who would volunteer to check a fifty page proof. Value judgements would enter; it would depend on what is at stake. A purported proof of the Riemann Hypothesis might attract more checkers than the sum of two excessively long integers. But one doesn't have to deal with fifty page proofs: most proofs in research papers are unchecked other than by the author. But then, most theorems are without issue: the last of a line of noble thought. They remain unchecked in the light of usage. They are loaded with errors.

If computing machines are employed either to check manipulation worked out by hand, or as has been done in some instances, to develop new theorems, the same remarks apply, but the probabilities may be altered. An interesting aspect of the problem of fidelity arises in programming. There are programs which are hundreds

of thousands of words and instructions long. Such programs are frequently written by batteries of programmers and the parts are spliced together. Now the problem is this: what in fact does the program do? Well, ask the programmers what it does. "My part works," says the first programmer over the phone from a laboratory 2000 miles away where he has just taken a new job. "So does mine," says the 2nd programmer who is still around but whose program is loaded with bugs that have not yet emerged. The third programmer: alas for flesh and blood, he died several months ago.

The program itself is the only complete description of what the program will do. This assumes that you know how the machine itself interprets a program—and this is not always the case. There may be no absolutely complete description of what the machine will do in a given instance. And all of this assumes that the machine treats its electronic symbols with perfect fidelity. To add to the indeterminacy, in a poorly designed computational system, the way the computer processes, my input may depend upon what my colleague down the hall is doing on his terminal. Cf. the concepts of fuzzy languages, algorithms, and environments. See, e.g., Zadeh [3]. This leads one to the pragmatic solution: run the program and you will see. You may learn that the performance is acceptable. In other cases you may not even be able to judge the quality of the output rationally. It may be a matter of faith.

Extremely long programs represent theorems of a kind. They may be far less trivial than some current frontier mathematics of conventional sort in terms of their distance from atomic symbolisms. But the problem is that we do not know and cannot know what the theorem says.

The upshot of this discussion is that the authenticity of a mathematical proof is not absolute, but only probabilistic. Proofs have attached to themselves lists of discoverers, sponsors, users, checkers, authenticators, rearrangers, generalizers, simplifiers, rediscoverers, swamis, communicants, and historians. These lists are all incorporated into the scholarly apparatus of publication and in the constant exposure that goes on the blackboard.

Proofs cannot be too long, else their probabilities go down and they baffle the checking process. To put it in another way: all really deep theorems are false (or at best unproved or unprovable). All true theorems are trivial.

A parallel with relativity theory can be made here. Newtonian mechanics grew up in a regime of low velocities and hence no relativity correction $(1-(v/v_c)^2)^{\frac{1}{2}}$ is necessary. Conventional (precomputer) mathematics grew up in a regime in which proof lengths were sufficiently low so that the fidelity could be considered absolute and the laws of information theory are irrelevant. It is also possible that mathematics might move into a period and into a corpus of material where the proof aspect ceases to have the classical significance and where one can live intimately with less than perfect fidelity.

6. On the observed incidence of error. What I have to say here is largely a collection of gossip. Since the subject is touchy, I shall begin at home.

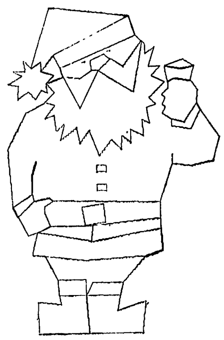


FIG. 3

A digitalized Santa is a mathematical object and its transformations are analogous to theorems. The aesthetic appeal of such theorems may have a different basis than that of classical mathematics. Less than perfect fidelity in processing is probably not very damaging.

The original printing of Davis, *Interpolation and Approximation*, contained at least 4 typewritten pages of errata. These range all the way from minor typos to errors of more mathematical substance. There is at least one bad proof and one theorem erroneously worded which if taken literally, is false. Davis and Rabinowitz, *Numerical Integration*, a smaller book whose galleys were proofread by both authors, has about a typewritten page of errors. One formula is just plain wrong. It was copied, without checking from the original author who worked it out wrong. Other errors are less easily alibied.

The original printing of *A Handbook of Mathematical Functions*, a thousand page compendium of formulas and tables which was put out by the National Bureau of Standards and which has sold more than 100,000 copies to date, contained more than several hundred errors. In the old days, when table making was a handcraft, some table makers felt that every entry in a table was a theorem (and so it is) and must be correct. Others took a relaxed, quality control attitude. One famous table maker used to put in errors deliberately so that he would be able to spot his work when others reproduced it without his permission.

I have before me a highly important book on advanced topics on analysis published about 15 years ago. After the book appeared, the author circulated to his friends an errata sheet of about 10 pages.

I have before me also the mimeographed 1925 notes of E. H. Moore of the University of Chicago on Hermitian matrices. One hundred eighty pages of notes are followed by 26 pages of errata.

There is a story to the effect that when B. O. Peirce's popular *A Table of Integrals* had just appeared, Professor Peirce offered a dollar to any student who discovered an error in it. Allowing an inflation rate of 3 or 4 to 1, I doubt whether any prudent author today would make a similar offer for his book. (D. E. Knuth has an open offer of this sort for his series of books on the art of computer programming.)

A recent issue of the *Notices* of the American Mathematical Society ran abstracts of about 130 papers: Five papers were listed as "Withdrawn". Presumably some of them had mistakes.

The *Mathematical Reviews* of December 1970, reports a paper entitled "The Decline and Fall of a Theorem of Zarankiewicz".

A past editor of the *Mathematical Reviews* once told me—somewhat in jest—that 50% of all mathematics papers printed are flawed.

A colleague reports refereeing a paper whose main theorem was invalid because the author spliced onto an erroneously stated theorem in a major reference book in topology. The words 'closed' and 'open' had inadvertently been interchanged in the reference.

There is a book entitled *Erreurs de Mathématiciens* by Maurice Lecat, published in 1935 in Brussels. This book contains more than 130 pages of errors committed by mathematicians of the first and second rank from antiquity to about 1900. There are parallel columns listing the mathematician, the place where his error occurs, the man who discovers the error and the place where the error is discussed. For example, J. J. Sylvester committed an error in "On the Relation between the Minor Determinant of Linearly Equivalent Quadratic Factors", Philos. Mag., (1851) pp. 295–305. This error was corrected by H. E. Baker in the *Collected Papers of Sylvester*, Vol. I, pp. 647–650.

In 1917 H. W. Turnbull calculated a system of 125 invariants of two quaternary quadratic forms. In 1929 Williamson found that three were reducible. In 1946, Turnbull himself found that five more were reducible, while in 1947, J. A. Todd found a further reducible one. Does it matter?

A mathematical error of international significance may occur every twenty years or so. By this I mean the conjunction of a mathematician of great reputation and a problem of great notoriety. Such a conjunction occurred around 1945 when H. Rademacher thought he had solved the Riemann Hypothesis. There was a report in *Time* magazine. Another instance was around 1860 when Kummer, following in the erroneous footsteps of Cauchy and Lamé, thought he had solved the Fermat Last Theorem.

8. Conclusions. Symbols and operations do not have a precise meaning, but only a probabilistic meaning.

A derivation of a theorem or a verification of a proof has only probabilistic validity. It makes no difference whether the instrument of derivation or verification is man or a machine. The probabilities may vary, but are roughly of the same order of magnitude when compared with cosmic probabilities.*

*E. Borel once suggested that the following chances constitute an unobservable event:

On the human scale:	1 chance in 10^6
On the terrestrial scale:	1 chance in 10^{15}
On the cosmic scale:	1 chance in 10^{50}
Absolute zero:	1 chance in 10^{500}

Mathematics has some of the aspects of an experimental science. We are saved from chaos by the stability of the universe which implies the repeatability of experiments and the self-correcting features of usage.

Mathematics has been Platonic for years. Does this rob it of a certain freedom and vitality which might be obtained by openly recognizing its probabilistic nature?

It is possible that a new type of mathematics might develop in which the "derivations" or the "processes" are so enormously long that the probabilistic nature of the result will be an integral feature of the subject.

References

1. N. Bourbaki, The architecture of mathematics, this MONTHLY, 57 (1950) 221–232.

2. E. Cerutti and P. J. Davis, FORMAC meets PAPPUS: Some observations on elementary analytic geometry by computer, this MONTHLY, 76 (1969) 895–905.

3. L. A. Zadeh, Fuzzy algorithms, Information and Control, 12(1963)94–102.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

COMPLETE ORTHONORMAL SYSTEMS IN PRE-HILBERT SPACES

MICHAEL GOLOMB, Purdue University

Summary. The usual concept of completeness of a system is of little use in a pre-Hilbert space since it does not imply linear density. It is replaced by "C-completeness", which has the same consequences in pre-Hilbert spaces that completeness has in Hilbert spaces.

1. Complete systems. Suppose $\{x_i\}$ is an orthonormal sequence in a separable inner product space S, that is

$$(x_i, x_j) = \delta_{ij} \qquad i,j = 1,2,\cdots. \tag{1}$$

Customarily one defines: $\{x_i\}$ is *complete* in S if there is no $x \neq 0$ in S which is orthogonal to the x_i, or, $\{x_i\}$ is *complete* if:

$$(x, x_i) = 0 \quad \text{for } i = 1,2,\cdots \text{ implies } x = 0. \tag{2}$$

It is well known that if the orthonormal sequence $\{x_i\}$ is *linearly dense* in S, that is, for each $x \in S$ and number $\varepsilon > 0$ there is an integer n and an element y in the linear span of $x_1, x_2, \cdots, x_n$ such that $\| x - y \| < \varepsilon$, then the system $\{x_i\}$ is complete.

In fact,

$$\left\| x - \sum_{i=1}^{n} (x,x_i)x_i \right\| \leqq \| x-y \| < \varepsilon; \tag{3}$$

hence if $(x,x_i) = 0$ $(i = 1,2,\cdots)$, then $\| x \| < \varepsilon$ for every $\varepsilon > 0$, so $x = 0$. There is a converse to this proposition: completeness of $\{x_i\}$ in S implies linear density in S, provided S is a complete space. For the partial sums of $\sum (x,x_i)x_i$ form a Cauchy sequence, hence $x - \sum (x,x_i)x_i$ is an element of the complete space S orthogonal to all x_i. Thus $x = \sum (x,x_i)x_i$, that is, $\| x - y \| < \varepsilon$ if $y = \sum_{i=1}^{n}(x,x_i)x_i$ with n sufficiently large.

However, if S is not complete, then the completeness of the sequence $\{x_i\}$ in the above sense does not imply linear density. We give a simple example, a special case of a more general one given in [**1**, p. 197]. Let $\boldsymbol{C}$ be the space of complex-valued functions defined and continuous in the interval $[-\pi,\pi]$, with the inner product

$$(x,y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t)\overline{y(t)}dt. \tag{4}$$

The functions e_k $(k = 0, \pm 1, \pm 2,\cdots)$, where $e_k(t) = e^{ikt}$, form an orthonormal system in $\boldsymbol{C}$ (not a sequence, but easily arranged as one). Let z be the discontinuous function for which $z(t) = 2$, $0 \leqq t \leqq \pi$, and $z(t) = 0$, $-\pi \leqq t < 0$, and set $\alpha_k = (z,e_k)$. Clearly $\alpha_0 = 1$. Introduce the system in $\boldsymbol{C}$

$$y_j = e_j - \alpha_j e_0 \qquad j = \pm 1, \pm 2,\cdots. \tag{5}$$

This system is not linearly dense in $\boldsymbol{C}$. In fact, $(z,y_j) = 0$ for $j = \pm 1, \pm 2,\cdots$, hence $\| z-y \| \geqq \| z \| = \sqrt{2}$ for any linear combination y of the y_j. Since there are functions in $\boldsymbol{C}$ arbitrarily close to z, we can find $z_1 \in \boldsymbol{C}$ such that $\| z_1 - y \| > 1$. On the other hand, if for some $x \in \boldsymbol{C}$ we have $(x,y_j) = 0$ $(j = \pm 1, \pm 2,\cdots)$, then either $(x,e_0) = 0$ and by (5) also $(x,e_k) = 0$ for $k = \pm 1, \pm 2,\cdots$, so $x = 0$; or $(x,e_0) = \gamma \neq 0$ and by (5), $(x - \gamma z, e_k) = 0$ for $k = 0, \pm 1, \pm 2,\cdots$, so $x = \gamma z$, which is impossible since $z \notin \boldsymbol{C}$. Thus the sequence $\{x_i\}$, that is obtained from $\{y_j\}$ by orthonormalization, is complete, but not linearly dense in $\boldsymbol{C}$.

The example shows that although the Fourier system $\{e_k\}$ is complete in $\boldsymbol{C}$, we cannot deduce that this system is linearly dense in $\boldsymbol{C}$, or that Parseval's equation $\| x \|^2 = \sum |(x,e_k)|^2$ holds for every $x \in \boldsymbol{C}$. We conclude that completeness in an incomplete space is rather pointless and propose the following stronger concept.

2. *C*-complete systems.

DEFINITION. The sequence $\{x_i\}$ in the inner product space S is *C-complete* if each Cauchy sequence $\{y_n\}$ in S for which

$$\lim_{n\to\infty} (y_n,x_i) = 0, \qquad i = 1,2,\cdots, \tag{6}$$

is a null sequence, i.e., $\lim \| y_n \| = 0$.

Although the definition applies to an arbitrary sequence $\{x_i\}$, we may restrict ourselves to orthonormal sequences, since C-completeness of $\{x_i\}$ is clearly equivalent to C-completeness of the sequence obtained from $\{x_i\}$ by orthonormalization.

We first observe that if S is complete, then $\{x_i\}$ is C-complete if and only if $\{x_i\}$ is complete. For, in this case, the Cauchy sequence $\{y_n\}$ has a limit $y \in S$, and (6) implies $(y, x_i) = 0$ for $i = 1, 2, \cdots$. Hence $y = 0$ if $\{x_i\}$ is complete. It is trivial that C-completeness implies completeness.

THEOREM. *The sequence $\{x_i\}$ is C-complete in S if and only if it is linearly dense in S.*

Proof. As remarked above, we may assume that $\{x_i\}$ is an orthonormal sequence. Suppose $\{x_i\}$ is C-complete and y is an arbitrary element of S. We set

$$y_n = y - \sum_{i=1}^{n} (y, x_i)x_i$$

and have $(y_n, x_i) = 0$ $(i = 1, \cdots, n)$, hence $\lim (y_n, x_i) = 0$ for $i = 1, 2, \cdots$. The sequence $\{y_n\}$ is Cauchy since

$$\| y_n - y_m \|^2 = \sum_{i=m+1}^{n} |(y, x_i)|^2 \quad \text{and} \quad \sum_{i=1}^{\infty} |(y, x_i)|^2$$

is convergent. It follows that $\lim \| y_n \| = 0$, and this proves the linear density of $\{x_i\}$.

Conversely, assume the sequence $\{x_i\}$ is linearly dense, and $\{y_n\}$ is a Cauchy sequence for which (6) holds. Given $\varepsilon > 0$, we determine N so large that $\| y_m - y_n \| < \varepsilon/3$ for $m, n \geqq N$. Since $\{x_i\}$ is linearly dense, the Parseval equation holds and

$$\begin{aligned} \| y_n - \sum_{i=1}^{j} (y_n, x_i)x_i \| &= \left\{ \sum_{i=j+1}^{\infty} |(y_n, x_i)|^2 \right\}^{\frac{1}{2}} \\ &\leqq \| y_n - y_N \| + \left\{ \sum_{i=j+1}^{\infty} |(y_N, x_i)|^2 \right\}^{\frac{1}{2}}. \end{aligned} \tag{7}$$

We choose $j = k$ sufficiently large that the last term in (7) is $< \varepsilon/3$. Then

$$\| y_n - \sum_{i=1}^{k} (y_n, x_i)x_i \| < \frac{2}{3}\varepsilon \qquad n > N. \tag{8}$$

Since (6) is assumed, the sum in (8) has norm $< \varepsilon/3$ if $n > N' > N$. Thus $\| y_n \| < \varepsilon$ for $n > N'$ and $\lim y_n = 0$, which proves the C-completeness of $\{x_i\}$.

Since linear density of the orthonormal sequence $\{x_i\}$ in S is trivially equivalent to the validity of the Parseval equation in S, we have the following result:

COROLLARY. *The orthonormal sequence $\{x_i\}$ is C-complete in S if and only if*

$$\| x \|^2 = \sum_{i=1}^{\infty} |(x, x_i)|^2 \tag{9}$$

for each $x \in S$.

3. The Fourier system. We now give an elementary proof of the C-completeness of the Fourier system $\{e_k\}$ which makes no use of the linear density of this system (Weierstrass Approximation Theorem, convergence of Fourier series to the function, etc.). The proof is essentially the same as the well-known one for completeness; see for example [**1**, p. 47] or [**2**, p. 11].

Suppose $\{y_n\}$ is a Cauchy sequence in the vector space E with inner product (4), where $C \subset E \subset L_2$, and

$$\lim_{n\to\infty} (y_n, e_k) = 0 \qquad k = 0, \pm 1, \pm 2, \cdots. \tag{10}$$

Let $y \in L_2$ be the limit of the sequence $\{y_n\}$. By (11),

$$(y, e_k) = 0 \qquad k = 0, \pm 1, \pm 2, \cdots. \tag{11}$$

For $Y(t) = \int_{-\pi}^{t} y(s)ds$ we have $Y(\pi) = 0$ by (10) for $k = 0$. Therefore integration by parts in (11) gives $(Y, e_k) = 0$ for $k = \pm 1, \pm 2, \cdots$. If we set $z = Y-(Y, e_0)e_0$, then

$$(z, e_k) = 0 \qquad k = 0, \pm 1, \pm 2, \cdots. \tag{12}$$

The function z is continuous, and if $z \neq 0$, say $z(t_0) = 2c > 0$, then there is an interval $I = [t_0 - \delta,\ t_0 + \delta]$ in which $z(t) \geqq c$. The function h defined by $h(t) = 1 + \cos(t - t_0) - \cos\delta$ is $\geqq 1$ in I and < 1 in the complement CI of I in $[-\pi, \pi]$. Now h^n is in the linear span of the e_k for each positive integer n, hence $(z, h^n) = 0$ by (12). But this is a contradiction since $\lim \int_{CI} zh^n = 0$, while $\int_I zh^n \geqq 2c\delta$ for each n. Therefore $z = 0$ and also $y = z' = 0$. Thus we have proved the system $\{e_k\}$ is C-complete in E.

The preceding proof makes use of the fact that the Cauchy sequence $\{y_n\}$ has a limit in L_2. We modify the proof so that no use is made of the Riesz-Fischer theorem, nor indeed of Lebesgue integration. We set

$$Y_n(t) = \frac{1}{2\pi}\int_{-\pi}^{t} y_n(s)ds \tag{13}$$

and observe that (10) for $k = 0$ gives $\lim Y_n(\pi) = 0$. Therefore integration by parts in (10) gives $\lim(Y_n, e_j) = 0$ for $j = \pm 1, \pm 2, \cdots$, and if we set $z_n = Y_n - (Y_n, e_0)e_0$, then also $\lim(z_n, e_k) = 0$ for $k = 0, \pm 1, \pm 2, \cdots$. The sequence $\{z_n\}$ converges uniformly in $[-\pi, \pi]$ since $|z_m(t) - z_n(t)| \leqq 2\| y_m - y_n \|$. For the continuous limit function z we have (12), and proceeding as above we conclude $z = 0$. It remains to show $\lim \| y_n \| = 0$.

Let $\varepsilon > 0$ be given and choose N so that $\| y_n - y_N \| < \frac{1}{2}\varepsilon$ for $n > N$. Also choose a step function u defined in $[-\pi, \pi]$ for which $\| y_N - u \| < \frac{1}{2}\varepsilon$, hence $\| y_n - u \| < \varepsilon$ for $n > N$. Then

$$\begin{aligned} \| y_n \|^2 &= \| y_n - u \|^2 - \| u \|^2 + 2\,\mathrm{Re}\,(y_n, u) \\ &\leqq \varepsilon^2 + 2\,\mathrm{Re}\,(y_n, u), \qquad n > N. \end{aligned} \tag{14}$$

Suppose the discontinuities of u are at the points t_i. Then (y_n, u) is a linear combinations of terms

$$\frac{1}{2\pi}\int_{t_i}^{t_{i+1}} y_n(s)ds = z_n(t_{i+1}) - z_n(t_i). \tag{15}$$

Hence $\lim(y_n, u) = 0$ and, by (14), $N_\varepsilon > N$ can be so chosen that $\| y_n \| < 2\varepsilon$ for $n > N_\varepsilon$. Therefore $\lim \| y_n \| = 0$ is proved.

From the C-completeness of the system $\{e_k\}$ we obtain a simple proof for a Fourier convergence theorem. Suppose f is a function of period 2π which has a piecewise continuous (more generally, a square-summable) derivative f'. Put

$$r_{m,n} = f - \sum_{k=-m}^{n} (f, e_k)e_k \qquad m, n = 0, 1, \cdots. \tag{16}$$

Since $f(\pi - 0) = f(-\pi + 0)$, integration by parts gives $(f', e_k)e_k = (f, e_k)e_k'$, hence $r'_{m,n} = f' - \sum_{k=-m}^{n}(f', e_k)e_k$. C-completeness of the system $\{e_k\}$ implies

$$\lim_{m,n\to\infty} \| r_{m,n} \| = 0, \quad \lim_{m,n\to\infty} \| r'_{m,n} \| = 0. \tag{17}$$

This, in connection with the trivial identity

$$tr_{m,n}(t) = \int_0^t r_{m,n}(s)ds + \int_0^t sr'_{m,n}(s)ds \tag{18}$$

gives $\lim_{m,n} r_{m,n}(t) = 0$ for each $t \neq 0$, but also for $t = 0$ since $r_{m,n}(2\pi) = r_{m,n}(0)$. This proves pointwise convergence of the Fourier series to f. Moreover,

$$\left| r_{m,n}(t) - r_{m,n}(-\pi) \right| = \left| \int_{-\pi}^{t} r'_{m,n}(u)du \right| \leqq \| r'_{m,n} \|, \tag{19}$$

and since $\lim r_{m,n}(-\pi) = 0$ and $\lim \| r'_{m,n} \| = 0$, (19) implies uniform convergence of the partial sums $\sum_{k=-m}^{n}(f, e_k)e_k$ to f.

References

1. S. Kaczmarz and H. Steinhaus, Theorie der Orthogonalreihen, 2nd ed., Chelsea, New York, 1951.

2. A. Zygmund, Trigonometric Series, Vol. 1, 2nd ed., Cambridge University Press, 1959.

HAAR INTEGRALS ON TOPOLOGICAL RINGS

JAMES T. SMITH, San Francisco State College

Let R be a locally compact topological ring with identity. Denote by R^+ its additive group, and by $R^\times$ the multiplicative group of its units, and assume that $R^\times$ is open in R. We shall give a simple method of constructing a Haar integral on $R^\times$ from a given Haar integral on R^+. This result complements, and its proof is suggested

by, the usual examples of Haar integrals. We then work out the particular example of the Haar integrals for the ring R of 2 by 2 real matrices.

Following Nachbin [2], we define a **right Haar integral** on a locally compact topological group G (notated additively) to be a nontrivial positive linear functional $\int$ on the vector space V of continuous real valued functions on G with compact support, such that for each $f \in V$ and $t \in G$,

$$\int f(x+t)\,dx = \int f(x)\,dx.$$

(Left Haar integrals may be defined and treated similarly.) Note that a right Haar integral on G always exists; further, if $\int_1$ and $\int_2$ are both right Haar integrals on G, then there exists a unique positive real number Δ such that $\int_1 = \Delta \int_2$. (We assume only these facts about the Haar integral, so it will be necessary to give a proof of a well-known property of the modulus function.)

Let R be a locally compact topological ring with identity, such that $R^\times$ is open in R. Then R^+ and $R^\times$ are locally compact topological groups under the topologies inherited from R. Let V^+ and $V^\times$ denote the vector spaces of continuous real valued functions with compact support on R^+ and $R^\times$, respectively. Let $\int^+$ be a Haar integral on R^+; we shall construct from $\int^+$ a right Haar integral $\int^\times$ on $R^\times$.

Let $t \in R^\times$. If $f \in V^+$, then the function that maps each $x \in R^+$ onto $f(xt)$ is also in V^+; define

$$\int_t^+ f(x)\,dx = \int^+ f(xt)\,dx. \tag{1}$$

THEOREM 1. *If $f \in V^+$, then the function that maps each $t \in R^\times$ onto $\int_t^+ f(x)\,dx$ is continuous.*

Proof. This results from the following inequality, which holds for all t and u in $R^\times$:

$$\left| \int_t^+ f(x)\,dx - \int_u^+ f(x)\,dx \right| \leqq \int^+ |f(xt) - f(xu)|\,dx.$$

THEOREM 2. *If $t \in R^\times$, then $\int_t^+$ is a Haar integral on R^+.*

Proof. Clearly, $\int_t^+$ is a nontrivial positive linear functional on V^+. Moreover, if $u \in R^+$, then

$$\int_t^+ f(x+u)\,dx = \int^+ f(xt+ut)\,dx = \int^+ f(xt)\,dx = \int_t^+ f(x)\,dx.$$

By Theorem 2, for each $t \in R^\times$ there exists a unique positive real number $\Delta(t)$ such that for each $f \in V^+$,

$$\int^+ f(x)\,dx = \Delta(t) \int_t^+ f(x)\,dx. \tag{2}$$

We call $\Delta(t)$ the **modulus of** t.

THEOREM 3. *The modulus function Δ is a continuous homomorphism from $R^\times$ to the multiplicative group of positive real numbers* [2, p. 77].

Proof. Continuity results from Theorem 1: use some $f \in V^+$ such that $\int^+ f(x)\,dx \neq 0$. For this f and any t and u in $R^\times$,

$$\Delta(tu)^{-1} \int^+ f(x)\,dx = \int_{tu}^+ f(x)\,dx = \int^+ f(xtu)\,dx = \int_u^+ f(xt)\,dx$$

$$= \Delta(u)^{-1} \int^+ f(xt)\,dx = \Delta(u)^{-1} \int_t^+ f(x)\,dx$$

$$= \Delta(u)^{-1}\Delta(t)^{-1} \int^+ f(x)\,dx.$$

Thus $\Delta(tu) = \Delta(t)\Delta(u)$.

If $f \in V^\times$, then the support of f excludes a neighborhood of 0 in R^+, hence we can extend the function f/Δ to a function in V^+ by setting $f(x)/\Delta(x) = 0$ for each $x \in R^+ - R^\times$. Then we define

$$\int^\times f(x)\,dx = \int^+ f(x)\Delta(x)^{-1}\,dx. \tag{3}$$

THEOREM 4. $\int^\times$ *is a right Haar integral on* $R^\times$.

Proof. Clearly, $\int^\times$ is a nontrivial positive linear functional on $V^\times$. Moreover, if $t \in R^\times$, then

$$\int^\times f(xt)\,dx = \int^+ f(xt)\Delta(x)^{-1}\,dx = \Delta(1/t)\int_{1/t}^+ f(xt)\Delta(x)^{-1}\,dx$$

$$= \Delta(t)^{-1} \int^+ f(x)\Delta(xt^{-1})^{-1}\,dx = \Delta(t)^{-1} \int^+ f(x)\Delta(x)^{-1}\Delta(t)\,dx$$

$$= \int^+ f(x)\Delta(x)^{-1}\,dx = \int^\times f(x)\,dx.$$

Example: the ring R of 2 by 2 real matrices. Here $R^\times$ is the group of invertible 2 by 2 real matrices, and $\int^+$ is the Lebesgue integral on real 4-space. We determine first the modulus function:

$$\Delta(t) = (\det t)^2. \tag{4}$$

This equation arises from a calculation with Jacobians: if $y = xt$ and the matrices x and y have entries x_{ij} and y_{ij}, respectively, then

$$\int^+ f(y)\,dy = \int^+ f(xt)\,J\,dx = J \int^+ f(xt)\,dx, \tag{5}$$

$$J = \frac{\partial(y_{11}, y_{12}, y_{21}, y_{22})}{\partial(x_{11}, x_{12}, x_{21}, x_{22})} = (\det t)^2. \tag{6}$$

Equation (4) then follows from (1), (2), (5), and (6). The Haar integral on $R^{\times}$ is given by Equations (3) and (4):

$$\int^{\times} f(x)\,dx = \int^{+} \frac{f(x)}{(\det x)^2}\,dx. \tag{7}$$

Note: This result generalizes theorems in Bourbaki [**1**, p. 33] and Weil [**3**, p. 89].

This work was supported by the National Research Council of Canada. The author acknowledges a suggestion of Professor Steven A. Gaal.

References

1. N. Bourbaki, Éléments de Mathématique, Livre VI: Intégration, Ch. 7, 8. Hermann, Paris, 1963.

2. L. Nachbin, The Haar Integral, Van Nostrand, Princeton, 1965.

3. A. Weil, Basic Number Theory, (Die Grundl. der math. Wissen., Bd. 144), Springer, Berlin, 1967.

GREGORY'S METHOD FOR NUMERICAL INTEGRATION

G. M. PHILLIPS, University of St. Andrews, Scotland

Recently Peters and Maley [**1**] obtained formulas of the form

$$h \sum_{i=0}^{n} f_i - h \sum_{j=0}^{m} A_j^m (f_j + f_{n-j}), \tag{1}$$

with $m \leqq n$, for approximating to the integral

$$\int_{x_0}^{x_n} f(x)\,dx.$$

In (1) the abscissas x_j are equally spaced, with $x_j = x_0 + jh$, $j = 0, 1, \cdots, n$, and f_j denotes $f(x_j)$. These integration rules are exact if $f \in \Pi_m$, the set of polynomials of degree not greater than m. In [**1**], for a given value of $m \leqq n$, each rule (1) is constructed by adding together contributions from the intervals $[x_0, x_j]$ and $[x_{n-j}, x_n]$ for $1 \leqq j \leqq m-1$ and $[x_j, x_{j+m}]$ for $0 \leqq j \leqq n-m$. Each contribution gives exact results for integrands $f \in \Pi_m$. This ingenious 'overlapping' method gives m times the required integral.

We shall show here that for m even, say $m = 2k$, the formulas (1) may be expressed in the form

$$h \sum_{i=0}^{n} f_i + h \sum_{i=0}^{2k} a_i(\Delta^i f_0 + (-1)^i \nabla^i f_n), \tag{2}$$

which is known as Gregory's integration formula; see [2], p. 135. The coefficients a_i are independent of k and n. We shall also derive a simple formula for calculating the a_i.

In (2) the forward difference operator Δ, depending on h, is defined by

$$\Delta f(x) = f(x+h) - f(x)$$

(see for example [2], p. 46), and higher order differences are defined recursively from

$$\Delta^{i+1} f(x) = \Delta(\Delta^i f(x)),$$

$i = 1, 2, \cdots$. We also define Δ^0 as the identity operator, $\Delta^0 f(x) = f(x)$. Similarly the backward difference operator ∇ is defined by

$$\nabla f(x) = f(x) - f(x-h),$$

and, again, higher order differences are defined recursively. It is easy to show ([2], p. 46) that

$$\Delta^i f_0 = \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} f_j \quad \text{and} \quad \nabla^i f_n = \sum_{j=0}^{i} (-1)^{-j} \binom{i}{j} f_{n-j}.$$

Therefore,

$$\Delta^i f_n + (-1)^i \nabla^i f_n = \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} (f_j + f_{n-j}).$$

By considering this formula for $i = 0, 1, \cdots$ it follows by induction that $f_j + f_{n-j}$ can be expressed as a linear combination of terms of the form $\Delta^i f_0 + (-1)^i \nabla^i f_n$. Thus, (1) can be expressed in the form (2) with A_j^{2k} replaced by new coefficients, say c_i^{2k}. We shall show later that these coefficients are, indeed, independent of k.

In (2) there is a good reason for terminating the second summation at even-order differences. For, as we shall see, this rule is exact if $f \in \Pi_{2k+1}$. If a further correction term

$$\lambda(\Delta^{2k+1} f_0 + (-1)^{2k+1} \nabla^{2k+1} f_n), \tag{3}$$

for any choice of λ, is added to the right side of (2), the resulting integration rule will still integrate exactly all integrands $f \in \Pi_{2k+1}$. This follows from the fact that (3) is zero, since the $(2k+1)$-th differences of a polynomial $\in \Pi_{2k+1}$ are constant. Thus there is not a unique formula involving $2k+1$ correction terms which integrates exactly all integrands $f \in \Pi_{2k+1}$. This is illustrated by the formula obtained by Peters and Maley [1] for the case $2k+1 = 3$, which is not the same as the corresponding Gregory formula.

To derive (2) directly, we begin with the Euler-Maclaurin summation formula. If $f \in \Pi_{2k+1}$, we have

$$\text{(4)}\qquad \int_{x_0}^{x} f(x)dx = h\sum_{i=0}^{n} f_i - \frac{h}{2}(f_0+f_n) - \sum_{j=1}^{k} \frac{B_{2j}}{(2_j)!} h^{2j-1}(f_n^{(2j-1)} - f_0^{(2j-1)}),$$

since all derivatives higher than the $(2k+1)$-th are zero and the $(2k+1)$-th derivatives are constant. The coefficients B_{2j} are the Bernoulli numbers as defined in [2] page 132, where the formula is derived. In (4) we replace derivatives at x_0 by differences, using the following relation given on page 79 of [2]:

$$\text{(5)}\qquad h^{2j-1}f_0^{(2j-1)} = (2j-1)! \sum_{i=2j-1}^{2k+1} \frac{S_i^{(2j-1)}}{i!}\Delta^i f_0,$$

where the $S_i^{(2j-1)}$ are the Stirling numbers of the first kind.

Similarly, we replace derivatives at x_n by backward differences, using

$$h^{2j-1}f_n^{(2j-1)} = (2j-1)! \sum_{i=2j-1}^{2k+1} (-1)^{i-1} \frac{S_i^{(2j-1)}}{i!} \nabla^i f_n.$$

Thus (4) becomes

$$\text{(6)}\qquad \begin{aligned} \int_{x_0}^{x_n} f(x)dx &= h\sum_{i=0}^{n} f_i - \frac{h}{2}(f_0+f_n) \\ &\quad - h\sum_{j=1}^{k} \frac{B_{2j}}{2j} \sum_{i=2j-1}^{2k+1} \frac{S_i^{(2j-1)}}{i!}[(-1)^{i-1}\nabla^i f_n - \Delta^i f_0], \end{aligned}$$

if $f \in \Pi_{2k+1}$. From the above definitions, it is easy to verify that $\nabla^i f_n = \Delta^i f_{n-i}$ and, as remarked above, the $(2k+1)$-th differences of a polynomial $f \in \Pi_{2k+1}$ are constant. We deduce that the upper limit in the third summation on the right of (6) may be replaced by $2k$, giving Gregory's formula

$$\text{(7)}\qquad \int_{x_0}^{x_n} f(x)dx = h\sum_{i=0}^{n} f_i + h\sum_{i=0}^{2k} a_i(\Delta^i f_0 + (-1)^i \nabla^i f_n),$$

$f \in \Pi_{2k+1}$. In (7),

$$\text{(8)}\qquad a_i = \begin{cases} -\frac{1}{2}, & i = 0 \\ \dfrac{1}{i!} \displaystyle\sum_{j=1}^{[\frac{i+1}{2}]} \frac{B_{2j}}{2j} S_i^{(2j-1)}, & i > 0, \end{cases}$$

where

$$\left[\frac{i+1}{2}\right]$$

denotes the integer part of $(i+1)/2$.

This establishes that the integration rule (7) holds with coefficients a_i which are independent of k and n. To show that the coefficients a_i are identical with the c_i^{2k} which resulted when (1) was written in the form (2), we consider the difference

between (2) as written and (2) with a_i replaced by c_i^{2k}. This difference is

$$\sum_{i=0}^{2k} (a_i - c_i^{2k})(\Delta^i f_0 + (-1)^i \nabla^i f_n) = 0,$$

for $n \geqq 2k$. An induction argument shows that

$$a_i - c_i^{2k} = 0, \quad 0 \leqq i \leqq 2k.$$

To see this, let $d_i = a_i - c_i^{2k}$. Putting $f = 1$, we deduce that $d_0 = 0$. Let us assume that $d_i = 0$, $0 \leqq i \leqq 2j - 2$. It follows from the equation above

$$d_{2j-1}(\Delta^{2j-1} f_0 - \nabla^{2j-1} f_n) + d_{2j}(\Delta^{2j} f_0 + \nabla^{2j} f_n) = 0 \tag{9}$$

for $f = x^{2j}$ and $f = x^{2j+1}$, since higher differences of these monomials vanish. It happens to be sufficient to consider only $f = x^{2j}$. First the identities

$$\Delta^{2j-1}(x^{2j}) = h^{2j-1}(2j)!x + \tfrac{1}{2}(2j-1)h^{2j}(2j)! \tag{10}$$

and

$$\Delta^{2j}(x^{2j}) = h^{2j}(2j)! \tag{11}$$

may be derived by inverting the formula (5) connecting derivatives and differences. Using these we obtain

$$h^{2j}(2j)![-(n-2j+1)d_{2j-1} + 2d_{2j}] = 0, \tag{12}$$

which has to hold for all $n \geqq 2k$. This implies that $d_{2j-1} = d_{2j} = 0$, and so by induction $c_i^{2k} = a_i$, $0 \leqq i \leqq 2k$. Thus the Peters and Maley formulas which are derived to integrate even-order polynomials exactly coincide with the corresponding Gregory formulas.

To obtain a simpler expression for a_i, we use the forward difference interpolation formula (see [2], p. 50). This is

$$f(x_0 + sh) = f_0 + \binom{s}{1}\Delta f_0 + \cdots + \binom{s}{2k+1}\Delta^{2k+1} f_0, \tag{13}$$

for $f \in \Pi_{2k+1}$. On integrating (13), we obtain

$$\int_{x_0}^{x_1} f(x)dx = hf_0 + h\sum_{i=0}^{2k} b_i \Delta^{i+1} f_0, \tag{14}$$

where

$$b_i = \int_0^1 \binom{s}{i+1} ds.$$

We now write down (7) with x_0 replaced by x_1 and add this to (14), after first writing $\Delta^{i+1} f_0 = \Delta^i f_1 - \Delta^i f_0$. This gives

$$(15)\quad \int_{x_0}^{x_n} f(x)dx = h\sum_{i=0}^{n} f_i + h\sum_{i=0}^{2k} a_i(\Delta^i f_0 + (-1)^i\nabla^i f_n) + h\sum_{i=0}^{2k}(a_i+b_i)(\Delta^i f_1 - \Delta^i f_0),$$

$f \in \Pi_{2k+1}$. Comparison of (15) and (7) shows that

$$(16)\qquad \sum_{i=0}^{2k} (a_i + b_i)\Delta^{i+1} f_0 = 0,$$

for $f \in \Pi_{2k+1}$. Putting $k = 0$ and $f(x) = x$, we see that $a_0 + b_0 = 0$. If we assume that $a_i + b_i = 0$ for $i = 0,1,\cdots,2k-2$, we may put $f(x) = x^{2k}$ and $f(x) = x^{2k+1}$ in turn in (16) to show that $a_i + b_i = 0$ for $i = 2k-1$ and $i = 2k$. It follows by induction that $a_i + b_i = 0$ for all i. That is,

$$(17)\qquad a_i = -\int_0^1 \binom{s}{i+1} ds,$$

which appears much simpler than (8).

The a_i are conveniently calculated from a recurrence formula, which will now be derived. For $|x| < 1$, we write

$$(18)\quad \sum_{i=0}^{\infty}\left[\int_0^1 \binom{s}{i} ds\right]x^i = \int_0^1\left[\sum_{i=0}^{\infty}\binom{s}{i}x^i\right]ds = \int_0^1 (1+x)^s ds = x/\log(1+x).$$

Thus

$$(19)\qquad x = \left(x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots\right)(1 - a_0x - a_1x^2 - \cdots).$$

On equating coefficients of x^{k+2}, we obtain the recurrence formula

$$(20)\qquad a_k - \frac{a_{k-1}}{2} + \cdots + (-1)^k\frac{a_0}{k+1} + (-1)^k\frac{1}{(k+2)} = 0,$$

for $k = 0,1,\cdots$. The first few values of the a_i, computed from (20), are $a_0 = -1/2$, $a_1 = 1/12$, $a_2 = -1/24$, $a_3 = 19/720$.

Acknowledgment. The author is indebted to Professor Anthony Ralston for most helpful suggestions concerning the presentation of the material.

References

1. G. O. Peters and C. E. Maley, Numerical integration over any number of equal intervals, this MONTHLY, 75 (1968) 741–744.
2. A. Ralston, A First Course in Numerical Analysis, McGraw-Hill, New York, 1965.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

POLYTOPES AND TRANSLATIVE EQUIDECOMPOSABILITY

H. HADWIGER, Universität Bern

Let E^n, $n \geqq 2$, be the Euclidean n-space with origin O, and $\boldsymbol{P}^n$ the class of all (convex) n-polytopes in E^n. Given $A \in \boldsymbol{P}^n$ we denote its (nonvoid) interior by A^0. T stands for the group of all translations of E^n, D for the group of all rotations which leave O fixed. For $0 < \lambda < \infty$ and $A \in \boldsymbol{P}^n$, set $\lambda A: = \{\lambda x: x \in A\}$.

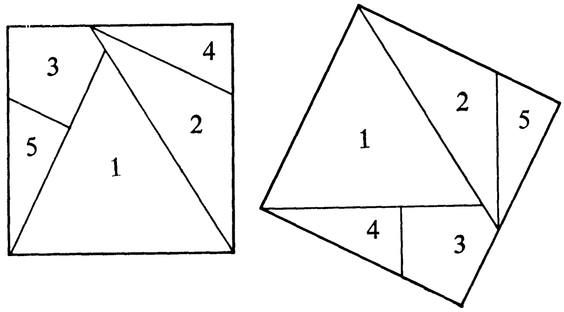
We write $A \overset{T}{\simeq} B$ if $A \in \boldsymbol{P}^n$ and $B \in \boldsymbol{P}^n$ are congruent by a translation, in other words, if there exists $\tau \in T$ such that $B = \tau A$. We say that $A, B \in \boldsymbol{P}^n$ are equidecomposable, with respect to the group of translations, and we write $A \overset{T}{\sim} B$, if there are families $(A_i)_{1 \leqq i \leqq r}$ and $(B_i)_{1 \leqq i \leqq r}$ of n-polytopes such that

$$A = \bigcup_1^r A_i,\ B = \bigcup_1^r B_i,\ (A_i \cap A_j)^0 = (B_i \cap B_j)^0 = \varnothing\ (i \neq j),$$

and $A_i \overset{T}{\simeq} B_i$, for all i. The notion of translative equidecomposability may be extended in a natural way to the class of all pairs of n-polyhedra (finite unions of n-polytopes) in E^n.

Let W be a fixed n-cube with edge-length 1. We consider four classes of polytopes:

I. An n-polytope belongs to $\boldsymbol{S}^n$ provided that it is centrally symmetric, and that all of its $(n-1)$-dimensional faces are centrally symmetric, too. Clearly $W \in \boldsymbol{S}^n$. By a theorem of Minkowski [**1**, p. 332], A lies in $\boldsymbol{S}^n$ if and only if the set of its $(n-1)$-faces is the disjoint union of two subsets $\{X_1, \cdots, X_r\}$ and $\{Y_1, \cdots, Y_r\}$ for which $X_i \overset{T}{\simeq} Y_i$, $1 \leqq i \leqq r$.



II. An n-polytope A belongs to $\boldsymbol{D}^n$ provided that $A \overset{T}{\sim} \delta A$, for all $\delta \in D$. It is well known [2] that W belongs to $\boldsymbol{D}^n$. The figure shows T-equivalent decompositions of two congruent squares.

III. An n-polytope A belongs to $\boldsymbol{W}^n$ provided that $A \overset{T}{\sim} \lambda W$, for some $\lambda > 0$. Trivially $W \in \boldsymbol{W}^n$.

IV. An n-polytope A belongs to $\boldsymbol{H}^n$ provided that A, and some finite union of homothets of A, are equidecomposable, with respect to T. In other words, $A \in \boldsymbol{H}^n$ if there exist a number $k \geqq 2$, reals $\lambda_i > 0$ and polytopes $A_i \in \boldsymbol{P}^n$ $(1 \leqq i \leqq k)$ such that $A \overset{T}{\sim} \bigcup_1^k A_i$, $(A_i \cap A_j)^0 = \varnothing$ $(i \neq j)$, and $A_i \overset{T}{\sim} \lambda_i A$. Considering a decomposition of W into $k = 2^n$ cubes W_i, such that $W_i \overset{T}{\simeq} (1/2)W$, we see that W belongs to $\boldsymbol{H}^n$. Our aim is to investigate the relations between these four classes of polytopes. First we remark that

$$\boldsymbol{S}^n \supset \boldsymbol{D}^n \supset \boldsymbol{W}^n \supset \boldsymbol{H}^n. \tag{1}$$

The proof of (1), which shall be omitted here, is based on a system of necessary conditions for the translative equidecomposability of two polytopes. These conditions are presented in [3] for the case $n = 3$. Our question is now, whether (1) may be replaced by the much stronger equality

$$\boldsymbol{S}^n = \boldsymbol{D}^n = \boldsymbol{W}^n = \boldsymbol{H}^n. \tag{2}$$

We suspect that (2) is true, for several reasons. The problem whether $\boldsymbol{S}^n = \boldsymbol{W}^n$, has been treated by E. Hertel (Jena) [6] for some time. The formal theory of polyhedral decomposition, as it has been developed by the author [4, p. 58 ff.] allows us to conclude $\boldsymbol{W}^n = \boldsymbol{H}^n$. In the case $n = 2$, our relation (2) holds almost trivially. $\boldsymbol{S}^3 = \boldsymbol{W}^3$ was proved some twenty years ago [5]. Recently H. R. Zobrist (Bern) [7] has shown $\boldsymbol{D}^3 = \boldsymbol{W}^3 = \boldsymbol{H}^3$. Thus our conjecture (2) is true at least for $n = 2$ and for $n = 3$. In order to establish it in all dimensions n, one would have to prove $\boldsymbol{H}^n \supset \boldsymbol{S}^n$.

References

1. B. Grünbaum, Convex Polytopes, Wiley, New York, 1967.

2. H. Hadwiger, Translative Zerlegungsgleichheit k-dimensionaler Parallelotope, Collectanea Math., 3 (1950) 3–15.

3. ———, Translative Zerlegungsgleichheit der Polyeder des gewöhnlichen Raumes, J. Reine Angew. Math., 233 (1968) 200–212.

4. ———, Vorlesungen über Inhalt, Oberfläche und Isoperimetrie, Springer, Berlin, 1957.

5. ———, Mittelpunktspolyeder und translative Zerlegungsgleichheit, Math. Nachr., 8 (1952) 53–58.

6. E. Hertel, Correspondence with the author, 1969/71.

7. H. R. Zobrist, Verschiedene Einzelstudien zur Zerlegungsgleichheit gewöhnlicher Polyeder, Diplomschrift, Bern, 1970.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

A FAMILIAR CONSTRUCTIBILITY CRITERION

KENNETH KALMANSON, Montclair State College.

The inclusion of some elementary field theory in an introductory abstract algebra course can pay handsome dividends in its dramatic applications, even if one stops considerably short of the "fundamental theorem of Galois theory." In particular, proof of the impossibility of certain ruler and compass constructions, such as the angle trisection, can be given on the basis of little more than basic theorems concerning field extensions. Many texts ([**1**]–[**4**] for example) develop the necessary condition:

(1) *A complex number z is constructible from the rational numbers Q only if* $[Q(z): Q] = 2^n$.

The status of the converse of (1), however, is not explicitly discussed in the texts cited above, even when, as in van der Waerden [**4**, page 185] a condition analogous to (1) that is both necessary and sufficient is developed. It may be stated as follows:

(2) *A complex number z is constructible from Q if and only if* $[K: Q] = 2^n$, *where K denotes the normal closure of $Q(z)$.*

Condition (2) presupposes more background, perhaps, than it is feasible to develop in an introductory undergraduate course. On the other hand, one does not want to leave the student with the mistaken idea that condition (1) is sufficient. One can correct this with an elementary counterexample as follows:

We show that at least one of the roots x_1, x_2, x_3, x_4 of the polynomial $f(x) \equiv x^4 + 4x + 2$ is not constructible. (Therefore none are.) By Eisenstein's criterion, $f(x)$ is irreducible over Q. Hence $[Q(x_i): Q] = 2^2$ for $i = 1, \cdots, 4$. Moreover, these roots can be found by Euler's method (cf. [**5**], pp. 121, 122). We have

$$x_1 = \sqrt{r_1} + \sqrt{r_2} + \sqrt{r_3}, \qquad x_3 = -\sqrt{r_1} + \sqrt{r_2} - \sqrt{r_3},$$
$$x_2 = \sqrt{r_1} - \sqrt{r_2} - \sqrt{r_3}, \qquad x_4 = -\sqrt{r_1} - \sqrt{r_2} + \sqrt{r_3},$$

where the r_j are the roots of the polynomial $g(t) \equiv t^3 - \frac{1}{2}t - \frac{1}{4}$, which is again irreducible (by Eisenstein applied to $8g(\frac{1}{2}s)$). Therefore, $[Q(r_1): Q] = 3$, and (1) implies that r_1 is not constructible. If x_1 and x_2 were constructible, then $r_1 = [\frac{1}{2}(x_1 + x_2)]^2$ would also be constructible, a contradiction.

Of course, if one knows (2) or, more specifically, that for any positive integer k

there exists an irreducible polynomial f_k over Q of degree k with Galois group the symmetric group on k letters (this result is proved in [4], section 61), then the status of the converse of (1) becomes apparent. Our method of proof, however, has the pedagogical advantages of being comparatively elementary and of motivating further discussion on topics such as solvability by radicals.

References

1. Iain T. Adamson, Introduction to Field Theory, Oliver and Boyd, ltd. Edinburgh, 1964, pp. 149–160.
2. W. E. Barnes, Introduction to Abstract Algebra, Heath, Boston, 1963, pp. 160–161.
3. Garrett Birkhoff and Saunders MacLane, A Survey of Modern Algebra, third edition, Macmillan, New York, 1970, pp. 379–380.
4. B. L. van der Waerden, Modern Algebra, vol. 1, (2nd ed.)., Frederick Ungar, New York, 1953, pp. 183–191.
5. W. S. Burnside and A. W. Panton, The Theory of Equations, Dublin University Press, 1924, pp. 121–122.

A CHARACTERIZATION OF COMPACT SUBSETS OF E^1

R. K. TAMAKI, California State College, Los Angeles

Our purpose is to prove the following result:

THEOREM. *Let $A \subset E^1$. Then the following are equivalent:*

(1) *A is compact.*

(2) *A has the fixed-set property, i.e., for every continuous $f: A \to A$ there is a nonempty subset $B \subset A$ such that $f(B) = B$.*

Note that the fixed-set property generalizes the well-known fixed-point property. It is a simple exercise [1, p. 252] to prove that every compact space has the fixed-set property, and that retracts of spaces having the fixed-set property also have the property. To prove the theorem we need the following simple lemmas:

LEMMA 1. *Each component of a subset $A \subset E^1$ having the fixed-set property is compact.*

Proof. If some component A_α is not compact it is either not closed in E^1 or it is unbounded, and in either case the student can easily construct a continuous f: $A \to A$ mapping A_α into A_α having no fixed-set.

LEMMA 2. *Let A be a set of positive reals such that A has closed components and contains arbitrarily small numbers. Then there exists a sequence of components $A_{\alpha_1}, A_{\alpha_2}, \cdots$ having 0 as a limit point, and such that*

$$\inf A_{\alpha_{i+1}} < \inf A_{\alpha_i}$$

for each $i = 1, 2, \cdots$. Furthermore, $B = \bigcup_i A_{\alpha_i}$ is a retract of A.

Proof. That the sequence can be selected is clear. Let A_{α_i} be any sequence satisfying the first two conditions. We shall show B is a retract of A. For every $k = 1, 2, \cdots$ select a positive x_k not in A such that $\sup A_{\alpha_{k+1}} < x_k < \inf A_{\alpha_k}$.

We now construct a retraction $\rho\colon A \to B$ with $\rho(x) = x$ for x in B and:

$$\rho(x) = \begin{cases} \sup A_{\alpha_1} & \text{if } \ x > \sup A_{\alpha_1} \\ \sup A_{\alpha_{k+1}} & \text{if } \ \sup A_{\alpha_{k+1}} < x < x_k \\ \inf A_{\alpha_k} & \text{if } \ x_k < x < \inf A_{\alpha_k} \end{cases}$$

ρ is continuous and hence is the desired retraction.

Proof of Theorem. We need only prove (2) implies (1), so assume $A \subset E^1$ has the fixed-set property. By Lemma 1 each component of A is compact. We conclude by showing that A is necessarily closed and bounded.

(i) A is closed. If not, take any limit point of A not in A, and without loss of generality, move it to the origin. Then use Lemma 2 to construct a retract of A not having the fixed-set property, an impossibility.

(ii) A is bounded. If A is unbounded, embedding E^1 (containing A) into $]0,1[$ will embed A as a nonclosed subset of E^1, contradicting (i).

Reference

1. J. Dugundji, Topology, Allyn and Bacon, Boston, 1966.

FINITE GEOMETRIES ON A TORUS

SISTER M. CORDIA EHRMANN, Villanova University

1. Finite geometry with metric. An interesting finite geometry, complete with metric, appears in [2] by Eves. It is an affine system in which the first 25 letters of the English alphabet are called **points.** The 30 sets of five letters that occur together in any row or any column of the following three blocks are called **lines.**

A	*B*	*C*	*D*	*E*	*A*	*I*	*L*	*T*	*W*	*A*	*X*	*Q*	*O*	*H*
F	*G*	*H*	*I*	*J*	*S*	*V*	*E*	*H*	*K*	*R*	*K*	*I*	*B*	*Y*
K	*L*	*M*	*N*	*O*	*G*	*O*	*R*	*U*	*D*	*J*	*C*	*U*	*S*	*L*
P	*Q*	*R*	*S*	*T*	*Y*	*C*	*F*	*N*	*Q*	*V*	*T*	*M*	*F*	*D*
U	*V*	*W*	*X*	*Y*	*M*	*P*	*X*	*B*	*J*	*N*	*G*	*E*	*W*	*P*

DEFINITION 1. By the **distance** Z_1Z_2 between two points Z_1 and Z_2 on a line p is meant the least number of steps along the line p from one point to the other, where the first letter of the line is considered as following the last letter of the line. (Thus on line $ABCDE$, distance $DB = 2$ and distance $AE = 1$.)

DEFINITION 2. A line q is **perpendicular** to a line p if there exist two points Z_1 and Z_2 on p such that $ZZ_1 = ZZ_2$ for each point Z on q. (Thus the line $AFKPU$ is perpendicular to the line $ABCDE$, since we may take $Z_1 = B$, $Z_2 = E$.)

Armed with the above identifications and definitions, it is possible to prove such non-trivial propositions as the following.

THEOREM 1. *The perpendicular bisectors of the three sides of a triangle are concurrent in a point.*

A little more development of the system allows the formulation and proof of the following theorem.

THEOREM 2. *The locus of the midpoints of a system of parallel chords of a parabola is a line perpendicular to the directrix of the parabola.*

2. Coordinatizing the geometry. The perhaps unexpected richness of so simple a geometry is further enhanced if we coordinatize the system after the method of Blumenthal [**1**].

$A\,(0,4)$	$B\,(1,4)$	$C\,(2,4)$	$D\,(3,4)$	$E\,(4,4)$
$F\,(0,3)$	$G\,(1,3)$	$H\,(2,3)$	$I\,(3,3)$	$J\,(4,3)$
$K(0,2)$	$L\,(1,2)$	$M(2,2)$	$N\,(3,2)$	$O\,(4,2)$
$P\,(0,1)$	$Q\,(1,1)$	$R\,(2,1)$	$S\,(3,1)$	$T\,(4,1)$
$U\,(0,0)$	$V\,(1,0)$	$W(2,0)$	$X\,(3,0)$	$Y\,(4,0)$

We can now define **slope** as the quotient of the difference of the ordinates and the difference of the abscissas of two points on the line. Since the coordinates are integers modulo 5, it may be considered desirable to use slopes that are also integers modulo 5. To reduce a fractional slope such as $\frac{2}{3}$ to an integer, we observe that 2 divided by 3 equals 4, modulo 5. Negative slopes can be disposed of with similar dispatch. It turns out that there are six parallel classes of lines, corresponding respectively to "no slope" and to slopes 0, 1, 2, 3, and 4. Each class contains five lines for a double-check total of 30 lines.

No slope:	$UPKFA$,	$VQLBG$,	$WRMHC$,	$XSNID$,	$YTOJE$.
Slope 0:	$UVWXY$,	$PQRST$,	$KLMNO$,	$FGHIJ$,	$ABCDE$.
Slope 1:	$UQMIE$,	$VRNJA$,	$WSOFB$,	$XTKGC$,	$YPLHD$.
Slope 2:	$ULCSJ$,	$VMDTF$,	$WNEPG$,	$XOAQH$,	$YKBRI$.
Slope 3:	$UGRDO$,	$VHSEK$,	$WITAL$,	$XJPBM$,	$YFQCN$.
Slope 4:	$UBHNT$,	$VCIOP$,	$WDJKQ$,	$XEFLR$,	$YAGMS$.

REMARK: In examining Eves' second and third blocks of 25 letters and comparing these with the coordinatized version of the first block, we see that he uses the fractional form of the slope to establish the order of points in a line.

3. A physical model. An article by Miller [3] suggests a physical model suitable

for the above system. The 30 lines with their total of 25 component points are arranged on a torus (two-dimensional torus in three-space), providing a nice visualization for the modular nature of the coordinate system. On the original coordinatized "square" model, one of its lines could lie on two or more parallel lines of the classical plane Euclidean superspace. On the torus, however, each line lies on a closed curve path wound about the torus. In fact, a line with slope k will wrap around the torus precisely k times, while passing laterally around the torus once.

4. A twenty-seven point system. In a similar vein, the author evolved a 27 point, 117 line system. (Picture a $3 \times 3 \times 3$ lattice-work cube.) The model emerged in the process of proving the *independence* of the following in Blumenthal's postulates for a finite affine system (stated here for a plane):

POSTULATE 4. *If p denotes any point, and m denotes any line, with p not an element of m, there is at most one line that contains p and has no point in common with m.*

It soon became obvious that this 27 point model would also (perhaps more appropriately) serve as the basis of a model for a postulate system of a Euclidean three-space, admitting of a finite interpretation. To enrich our 27 point system with planes, we simply assume that any three non-collinear points determine a plane.

The computation of the number of planes (or lines) can be accomplished by the use of combinatorial formulas. By this method it was determined that there are 39 planes and 117 lines in our model. These counts were subsequently confirmed by special formulas in Wylie [**4**].

5. Direction numbers. A more geometric approach to the counting of lines and planes was developed by the author via the improvisation of *direction numbers.* With this end in view, we first coordinatize the system. Let $\Sigma = \{0,1,2\}$. Assign the 27 ordered triples of $\Sigma \times \Sigma \times \Sigma$ to the 27 points, positioned as lattice points of classical Euclidean three-space. Now take a fresh look at the 27 ordered triples, this time as possible direction numbers. We summarily discard $(0,0,0)$ as literally "getting us nowhere," and note that the remaining triples can be identified in pairs, modulo 3. For example, $(2,0,1) = 2(1,0,2)$. This reduces the system to 13 "independent" ordered triples. Laborious rechecking by direct methods confirms the conjecture that there are indeed exactly 13 parallel classes of lines, each class containing nine lines.

The same 13 ordered triples can serve as "orthogonal" direction numbers for the respective planes. Since each plane contains nine points and there is a total of 27 points, each parallel class of planes must contain exactly three planes. This results in a total of 39 planes which once more checks with the answer obtained by more tedious procedures.

6. On a torus again. In order to surmount the apparent problem of a line or

plane lying on more than one line or plane of the classical Euclidean superspace of $\Sigma \times \Sigma \times \Sigma$, we may find it helpful to arrange the 27 points on a three-dimensional torus in four-space. That is, let each of the original three dimensions loop back upon itself, thus circumventing the difficulty.

7. Defining a metric. Something of a surprise may occur when we try to define a "reasonable" metric on our 27 point system. It turns out that we can consider the "most natural" metric on this finite geometry to be the *trivial metric*, whereby every two distinct points have a distance of exactly one unit between them. This can be rendered more credible in the classroom by looking at the nine point 12 line finite Euclidean two-space (Young's geometry) on a two-dimensional torus in three-space.

8. Classroom problem. Let $\Sigma = \{0, 1, 2, 3, 4\}$. How would one go about defining a nontrivial metric on $\Sigma \times \Sigma \times \Sigma$, such that the metric preserves the truth of many of the key theorems of classical Euclidean three-space?

References

1. L. M. Blumenthal, A Modern View of Geometry, Freeman, San Francisco, 1961, pp. 49–50.
2. H. Eves, A Survey of Geometry, Vol. 1, Allyn and Bacon, Boston, 1963, pp. 432–433.
3. W. A. Miller, A construction of and physical model for finite Euclidean and projective geometries, Math. Teacher, No. 4, 63 (1963) 301–306.
4. C. R. Wylie, Jr., Foundations of Geometry, McGraw-Hill, New York, 1964, pp. 50–51.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

A LABORATORY AND COMPUTER BASED APPROACH TO CALCULUS

SOLOMON GARFUNKEL, University of Connecticut

Introduction. The Education Research Center at M. I. T. is undertaking an experiment in science education called the Unified Science Study Program (USSP). This program is offered to one hundred freshmen and sophomores from M. I. T., Tufts, North Shore Community College, and the University of Massachusetts (at Boston). The underlying rationale of the program is that a student engages in a project, or series of projects, learning in response to the need for knowledge arising

ing. One survey respondent suggested that having an MAA representative at each two-year college in the section might provide greater input from the two-year colleges.

Other means of involvement. Visiting lecture programs for the two-year colleges are just beginning. More and more faculty from the four-year institutions are being seen on the two-year college campuses, and they are finding out that more innovations are taking place in the two-year colleges than in the tradition-bound four-year institutions.

A survey on two-year college faculty participation in mathematics organizations was conducted in a geographically large section. (See *The Two-Year College Mathematics Journal*, Vol. 2, No. 1, spring 1971, pp. 53–57.) This section is trying to improve its annual meetings to take account of the suggestions obtained by the survey. In another section, financial support for a speaker for the two-year college portion of a regional mathematics conference was provided by the section. Quite often, activities of a section hinge on the initiative in one institution and more particularly on one individual.

The officers of the Mathematical Association of America have said that they welcome suggestions from anyone on what else the MAA or its sections might do to be of greater service to the two-year college teachers of mathematics and that all such suggestions will be given serious consideration and implemented wherever possible.

THE U.S.A. MATHEMATICAL OLYMPIAD

Nura D. Turner, State University of New York at Albany

The first U.S.A. Mathematical Olympiad [1], a new activity of the Mathematical Association of America, will be held Tuesday, May 9, 1972.

Who will participate in this first U.S.A. Mathematical Olympiad? Just as the British use the Annual High School Mathematics Competition [2] as the qualifying round for their British Mathematical Olympiad, so we shall use that competition for our Olympiad. For this first one, invitations for participation will be extended to the approximately 100 top-ranking students who participated in the 1972 Annual High School Mathematics Competition (AHSMC). Additional participants will include students from States not involved in the AHSMC; they will be selected from participants in the comparable competitions in those States and on some proportional basis.

Quality and adequate time for reflection on and response to the problems will be emphasized. Students will sit in their own schools for the examination which will be composed of five essay-type questions requiring mathematical mature thinking for solution. Students will have three hours to think through the problems, organize proofs, and possibly, come forth with unique solutions.

Provision has been made for the expediting of grading that will include uniformity of grading. Each student will present solutions for each of the five problems in different booklets. Solutions for all No. 1 problems will be mailed to one member of the grading committee. Solutions for each of the other problems will be similarly handled. Results should be known by early in June.

According to plan, the top-ranking students, possibly eight or so, will be brought together during the summer to be honored in a suitably dignified ceremony.

This higher level testing in secondary school mathematics hopefully will have a far-reaching effect upon the mathematical atmosphere of our high schools. The experience with subjective-type testing will help fulfill an existing need in our country. Thought provoking questions will provide stimulation and challenge for our secondary school students highly talented in mathematical ability. For the first time we shall be operating in secondary school mathematics testing on a level with the British and Eastern European countries. And we can look forward to identification by this competition of students with creative minds.

Any questions can be directed to Professor Samuel L. Greitzer, Rutgers, The State University, Newark, New Jersey 07102.

References

1. Nura D. Turner, Why can't we have a USA Mathematical Olympiad? this MONTHLY, 78 (1971) 192–195.

2. Though the identical competition, it is known in Great Britain as the "National Mathematical Contest".

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before June 30, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2293 [1971, 405]. *Proposed by Erwin Just, Bronx Community College*

Does there exist an infinite set of primes, S, such that whenever $p \in S$ and $q \in S$, we have $(\frac{1}{2}(p-1), \frac{1}{2}(q-1)) = 1$, $(p, q-1) = 1$ and $(p-1, q) = 1$?

Editorial Note. The editors are embarrassed to announce that the files on this and problem E 2294 (below) seem to be lost. We request all solvers to resubmit their solutions.

E 2294 [1971, 405]. *Proposed by Douglas Lind, Stanford University*

For what n does the regular n-simplex of side 1 have rational height?

E 2343. *Proposed by G. A. Heuer, Concordia College*

According to a well-known theorem of analysis, a series of real numbers is unconditionally convergent (i.e., $\Sigma a_{\phi(n)} = \Sigma a_n$ for every permutation ϕ of the positive integers) if and only if it is absolutely convergent. Certain kinds of rearrangements, however, will leave the sum of an arbitrary convergent series unaltered. (A) Prove that if $\phi(n) - n$ is bounded, and Σa_n is any convergent series, then $\Sigma a_{\phi(n)} = \Sigma a_n$. (B) Prove or disprove: If $\phi(n) - n$ is unbounded, then there is a series Σa_n for which $\Sigma a_{\phi(n)} \neq \Sigma a_n$.

E 2344. *Proposed by Jordi Dou, Barcelona, Spain*

Consider a square array of red dots and blue dots with 50 rows and 50 columns. Whenever two dots of the same color are adjacent in the same row or column connect them with a segment of that color; if they are adjacent of different color, connect them with a black segment. There are 1269 red dots, among them 99 on the border, none of them at the corners. There are 1035 black segments. Find the number of red segments and the number of blue segments.

E 2345*. *Proposed by E. S. Langford, University of Maine*

Let S be any nonempty compact subset of the plane. A sequence $\{P_n\}$ of points of S has the following property:

$$d(P_n, P_{n+1}) = \max\{d(P_n, P): P \in S\}.$$

Let $d_n = d(P_n, P_{n+1})$. Then obviously $d_1 \leqq d_2 \leqq \cdots \leqq \delta$, where δ is the diameter of S. Let $d = \lim d_n$. (a) Is it possible that $d < \delta$? (b) Is it possible that the sequence $\{d_n\}$ is strictly increasing? (c) Is it possible that $\{d_n\}$ is strictly increasing and, in addition, that $d < \delta$?

E 2346. *Proposed by Louis Shapiro, Howard University*

Say that a group has *small centralizers* if every non-identity element commutes only with its inverse, itself, and the identity. Characterize all groups with small centralizers.

E 2347. *Proposed by L. Carlitz, Duke University*

Let P denote a point in the interior of the triangle ABC. Let α, β, γ denote the angles of ABC. Let R_1, R_2, R_3 denote the distances from P to the vertices of ABC, and let r_1, r_2, r_3 denote the distances of P from the sides of ABC. Show that

$$R_1^2 \sin^2\alpha + R_2^2 \sin^2\beta + R_3^2 \sin^2\gamma \leqq 3(r_1^2 + r_2^2 + r_3^2)$$

with equality if and only if P is the symmedian point of ABC.

E 2348. *Proposed by L. Carlitz, Duke University*

Let P be a point in the interior of the triangle ABC. Let R_1, R_2, R_3 denote the distances from P to the vertices of ABC and let r_1, r_2, r_3 denote the perpendicular distances from P to the sides of ABC. Show that

$$\Sigma R_1(r_2 + r_3) \geqq \Sigma(r_1 + r_2)(r_1 + r_3), \tag{1}$$

$$\Sigma(R_1 + R_2)(R_1 + R_3) \geqq 4\,\Sigma(r_1 + r_2)(r_1 + r_3), \tag{2}$$

with equality if and only if ABC is equilateral and P is its center.

SOLUTIONS OF ELEMENTARY PROBLEMS

Connected Graphs and Frequency Partitions

E 2277 [1971, 195]. *Proposed by Phyllis Chinn, Towson State College, Maryland*

A *graph* is a finite collection of points, and lines between them, where each line has two distinct endpoints and no two lines have the same pair of endpoints. The *degree* (or valency) of a point is the number of edges to which it belongs. The *partition* associated with a graph is the sequence of degrees of points in the graph. A *frequency partition*, which is a partition of the order of a graph, can be formed by recording the frequency with which each degree is assumed.

Prove that for any partition of an integer p, except $p = 1 + 1 + \cdots + 1$, there is a connected graph of order p having the given partition as its frequency partition.

Editorial Note. Without notifying the Problem Department, the proposer submitted her work elsewhere and her solution has appeared in *Recent Trends in Graph Theory*, Springer Verlag, 1971, pp. 69–71.

Also solved by Neal Felsinger, B. R. Myers, J. A. Roberts, H. S. Sun, J. J. Tattersall, the proposer, and an unknown solver.

Prime Divisors of Polynomials

E 2287 [1971, 298]. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

If P is a nonconstant polynomial with integral coefficients and k is any integer, must there exist an integer m for which there are at least k distinct prime divisors of $P(m)$?

Editor's comment: The answer is yes, and by coincidence a proof appeared in the same issue as did E 2287. See Irving Gerst and John Brillhart, *On the prime divisors of polynomials*, this MONTHLY 78 (1971), 250–266, Theorem 1, p. 253. This was noted by the following sharp-eyed readers: Anders Bager (Denmark), R. J. Dickson, Neal Felsinger, Ella Mae McIntyre, and Joy Rietmulder. Solutions were submitted by Frederick Carty, Don Coppersmith, G. A. Heuer & C. V. Heuer, Myron Hlynka, Harry Lass, Simeon Reich (Israel), St. Olaf College Students, Allen Stenger, E. W. Trost (Switzerland), K. L. Yocom, and the proposers.

There is no Inconsistency without "Not"

E 2288 [1971, 298]. *Proposed by John Corcoran, State University of New York at Buffalo*

Let L be the set of sentences of any predicate logic whose logical symbols are: the universal and existential quantifiers, identity, negation, conjunction, disjunction, implication. Does every inconsistent set of sentences from L contain at least one negation sign?

Solution by the proposer. Look at the set L^* of sentences devoid of negation signs. Notice that it is constructed recursively by the usual rules excluding the negation rule. Consider the interpretation (or model) I whose domain D is a singleton and which assigns D^n to each n-ary predicate and assigns the unique n-ary function from D to D to each n-ary function symbol. All atomic formulas are true for all values of variables in D (under I). All negation-free truth-functional combinations of true formulas are true. All quantifications of formulas true in I are true. Thus all sentences in L^* are simultaneously true in I, so that L^* is consistent. Thus every subset of L^* is consistent, and the answer to the problem is "yes."

Also solved by Kenneth Bowen and Neal Felsinger.

A Polynomial Identity

E 2290 [1971, 405]. *Proposed by E. H. Davis, Kansas State College at Pittsburg*

Describe all polynomials, $p(x,y)$, with real coefficients such that

$$p(x,y) = p(x+1,\ y+1).$$

Solution by William Franke, Abdolhamid Mohtadi, Jean Oyster, and Thomas Pickett, Students at Miami University (Ohio). Every such polynomial must be of the following form:

$$p(x,y) = \sum_{i=0}^{n} a_i(x-y)^i,$$

where each a_i is real. (Conversely, every such polynomial must satisfy $p(x,y) = p(x+1,\ y+1)$ for all x, y.)

Consider a polynomial $p(x,y)$ which satisfies the condition of the problem. Make the change of variables $x = u + v$ and $y = u - v$ so that $p(x,y)$ becomes a polynomial $f(u,v)$ in the variables $u = \frac{1}{2}(x+y)$, $v = \frac{1}{2}(x-y)$. Then $f(u,v) = f(u+1,v)$ and in fact $f(u+n,v) = f(u,v)$ for every integer n. If (a,b) is any point in the plane then the polynomial in one variable $g(u) = f(u,b) - f(a,b)$ has infinitely many zeros, so that $f(u,b) = f(a,b)$ for all u. In general, $f(u,v) = f(0,v)$ for all v, so that $p(x,y) = p(0, \frac{1}{2}(x-y))$, as was to be shown.

Also solved by sixty-seven other readers.

A number of solvers (and the proposer) note that their solutions hold for polynomials over any field of characteristic zero. Several generalizations were made: Jerzy Tiuryn (Poland) characterizes polynomials such that $p(ax + c, ay + c) = p(bx + d, by + d)$ for nonzero a and b and arbitrary c and d, and R. L. Snyder considers the problem for polynomials over an arbitrary field. The proposer remarks that he encountered the problem in generalizing the construction of non-planar nearfields given by J. Zemmer in *Mathematics Student*, 31 (1964), 145–150.

An Identity

E 2291 [1971, 405]. *Proposed by Barry Wolk, University of Manitoba*

If $\Sigma_0 f(n)$ means $f(1)+f(3)+f(5)+\cdots$, show that for all real x

$$\left|\sum_0 n^{-2}\cos(nx)\right| = \sum_0 n^{-2}\cos^2(nx).$$

Solution by Agnes Briggs, undergraduate, University of Pittsburgh. Let $f(x)=\frac{1}{4}\pi(\frac{1}{2}\pi - |x|)$ for $-\pi \leqq x \leqq \pi$. If $F(x)$ is the 2π-periodic extension of $f(x)$, then

$$F(x) = \sum_0 n^{-2}\cos nx.$$

Set $g(x)=f(x)$ for $-\frac{1}{2}\pi \leqq x \leqq \frac{1}{2}\pi$. The π-periodic extension of $g(x)$ is

$$\begin{aligned} |F(x)| &= \frac{\pi^2}{16} + \frac{1}{2}\sum_0 n^{-2}\cos 2nx \\ &= \frac{\pi^2}{16} + \sum_0 n^{-2}\cos^2 nx - \frac{1}{2}\sum_0 n^{-2} = \sum_0 n^{-2}\cos^2 nx, \end{aligned}$$

and the identity follows.

Also solved by thirty-seven other readers.

Direct Sums and Products of Infinitely Many Copies of the Integers

E 2292 [1971, 405]. *Proposed by Stephen Maurer, Phillips Exeter Academy*

Let S be the direct sum $\Sigma_{i\in I}Z_i$, and P the direct product $\prod_{i\in I}Z_i$, where each Z_i is a copy of the additive group of integers and I is an infinite set. Is the natural image of S in P a direct summand?

Solution by D. Ž. Djoković, University of Waterloo. The direct sum S is not a direct summand of P because P/S has a nonzero element which is divisible by every positive integer whereas P does not. To show this, we can assume that $N \subseteq I$, where N denotes the set of natural numbers. If we define $f \in P$ by $f(i) = i!$ if $i \in N$ and $f(i) = 0$ otherwise, then the element $f + S \in P/S$ has the required properties.

Also solved by Dennis Bertholf, Samuel Cox, Jr., Neal Felsinger, W. Margolis, Joel Spencer, and the proposer & Robert MacPherson.

Spencer comments that this problem was floating around Princeton a few years ago, and Margolis

notes that the problem is essentially stated and solved in E. Schenkman, *Group Theory*, Exercise II. 5. g. Cox generalizes by replacing Z by an arbitrary ring R with identity, and shows that if S is a direct summand of P as left R-modules, then R has the descending chain condition on finitely generated right ideals. He remarks further that if R is right coherent and has the DCC on finite right ideals, then S is a direct summand of P for each choice of I.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before June 30, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

5842. *Proposed by B. B. Winter, Eugene, Oregon*

Let T be a linear (not necessarily continuous) map of a Hilbert space H to itself. Suppose there exists a subset S such that $Tx \in S$ and $x - Tx \in S^{\perp}$ for all $x \in H$. Show that S is a closed linear subspace and that T is the (necessarily continuous) orthogonal projection of H onto S.

5843. *Proposed by N. P. Callas, Office of Scientific Research, U.S. Air Force*

Show that if $\sigma(x) \geqq 0$ satisfies the nonlinear differential inequality

$$\sigma'(x) + b(x)\sigma(x) \leqq f(x)[\sigma(x)]^{\alpha},$$

where $\sigma(a) = c$ and $0 \leqq \alpha < 1$, then $\sigma(x) \leqq$

$$\exp\left(-\int_a^x (1-\alpha)b(\tau)d\tau\right)\left[\int_a^x (1-\alpha)f(\tau)\exp\left(\int_a^\tau (1-\alpha)b(t)dt\right)d\tau + c^{1-\alpha}\right]^{1/(1-\alpha)}.$$

5844. *Proposed by L.-S. Hahn, University of New Mexico*

Construct a function defined everywhere in the plane which is nowhere continuous and yet is continuous in each variable separately, or prove such a function does not exist.

5845.* *Proposed by J. A. Johnson, Oklahoma State University*

Let X be an uncountable set and $\mathscr{A}$ the smallest σ-algebra of subsets of $X \times X$ containing all sets of the form $A \times B$ where $A \subset X$, $B \subset X$. Does $\mathscr{A}$ contain all subsets of $X \times X$?

5846. *Proposed by H. Kestelman, University College, London, England*

If $f \in L(0,\infty)$ and $I(\lambda)$ is, for each positive λ, a subinterval of $(0,\infty)$, then $\lim_{\lambda\to\infty}\int_{I(\lambda)} f(t)\cos\lambda t\, dt = 0$. If $I(\lambda)$ is assumed only to be the union of a finite set of intervals, the result is false.

5847. *Proposed by Joe Beasley, Prairie View A. & M. College*

X is a complete metric space and $T{:}X \to X$ is a function with the following conditions:

(1) There is a sequence $\{x_n\} \in X$ such that $d(x_n, T(x_n)) \to 0$.

(2) $t: X \to R$ defined by $t(x) = d(x, T(x))$ is lower semicontinuous.

(3) $d(T(x), T(y)) \leqq ad(x, T(x)) + bd(y, T(y)) + cd(x, y)$, where a, b, c are positive numbers and $c < 1$.

Show that (A) T has a unique fixed point, and (B) none of conditions (1), (2) or (3) can be omitted.

SOLUTIONS OF ADVANCED PROBLEMS

A Formula in GF(2^n)

5746 [1970, 774]. *Proposed by Leonard Carlitz, Duke University*

Let $GF(2^n)$ denote the finite field of order 2^n. For $a \in GF(2^n)$ put

$$e(a) = (-1)^{t(a)}, \ t(a) = a + a^2 + a^{2^2} + \cdots + a^{2^{n-1}},$$

$$S(a) = \sum_{x\,y\,z} e\left\{x + y + z + \frac{a}{yz + zx + xy}\right\},$$

where the summation is over all $x, y, z \in GF(2^n)$ such that $yz + zx + xy \neq 0$. Show that

$$S(a) = (-1)^n 2^n \sum_{x \neq 0} e(x + ax'), \qquad (xx' = 1).$$

I. *Solution by A. A. Jagers, Enschede, Holland.* Note first that, since $t(u)$ is the trace of u relative to the prime field P contained in $GF(2^n)$, $t(u) \in P$ for all u so that it is not difficult to decide what is meant by $(-1)^{t(u)}$. Put

$$\alpha = \sum_{yz+zx+xy=0} (-1)^{t(x+y+z)}, \ q_k = \left|\{(x, y, z) \,|\, yz + zx + xy = 0, \ x + y + z = k\}\right|$$

and $q = \Sigma_k q_k$. Then $\alpha = q_0 + \Sigma_{k \neq 0}(-1)^{t(k)} q_k$. Noting that $t(p^{\frac{1}{2}}) = t(p)$, it follows by substituting $(x, y, z) = (x' + p^{\frac{1}{2}}, y' + p^{\frac{1}{2}}, z' + p^{\frac{1}{2}})$ that

$$\sum_{yz+zx+xy=p} (-1)^{t(x+y+z)} = \alpha(-1)^{t(p)}.$$

Consequently

$$S(a) = \sum_{p \neq 0} \sum_{yz+zx+xy=p} (-1)^{t(x+y+z)}(-1)^{t(a/(yz+zx+xy))}$$
$$= \alpha \sum_{p \neq 0} e(p + ap^{-1}).$$

Now t is a P-linear functional and, because of the separability of the extension $GF(2^n)$ of P, $t \neq 0$. This implies that $\left|\{k \,|\, t(k) = 0\}\right| = \left|\{k \,|\, t(k) = 1\}\right|$ and thus that $\Sigma_k(-1)^{t(k)} = 0$. Moreover, the substitution $(x, y, z) = k(x', y', z')$ shows that $q_k = q_1$ for all $k \neq 0$. Hence $\alpha = q_0 - q_1$ and $q = q_0 + (2^n - 1)q_1$. Since

$$\left|\{(x, y, z) \,|\, yz + zx + xy = k\}\right| = q,$$

as follows from the substitution $(x,y,z) = (x' + k^{\frac{1}{2}},\ y' + k^{\frac{1}{2}},\ z' + k^{\frac{1}{2}})$, one has $2^n q = 2^{3n}$ and thus $(2^n - 1)\alpha = 2^n(q_0 - 2^n)$. Now $q_0 = \left|\{(x,y) \mid x^2 + y^2 = xy\}\right| = 1 + \left|\{(x,y) \mid x \neq 0,\ y \neq 0,\ u = xy^{-1},\ u + u' = 1\}\right| = 1 + (2^n - 1)\left|\{u \mid u^3 = 1, u \neq 1\}\right| = 1 + (2^n - 1)\ (1 + (-1)^n)$, since the multiplicative group of $GF(2^n)$ is cyclic with order $2^n - 1$ and since $2^n - 1 \equiv 0 \bmod 3$ if and only if n is even. Hence $\alpha = (-1)^n 2^n$ and $S(a) = (-1)^n 2^n \sum_{p \neq 0} e(p + ap^{-1})$, as desired.

II. *Remark by the proposer.* For $q = p^n$, p prime, $n \geqq 1$, $a \in F = GF(q)$, put

$$e(a) = e^{2\pi i t(a)/p},\ t(a) = a + a^p + \cdots + a^{p^{n-1}}.$$

$$K_1(a) = \sum_{\substack{x \in F \\ x \neq 0}} e(x + ax'),$$

$$K_2(a) = \sum_{\substack{x,y \in F \\ x \neq 0, y \neq 0}} e(x + y + ax'y'),$$

where $xx' = yy' = 1$. Also put $S(Q,L) = \sum_{(x)} e\{L(x) + (Q(x))^{-1}\}$, where $L(x)$ is a linear form and $Q(x)$ a quadratic form in $x_1, x_2, \cdots, x_s$ with coefficients in F and the summation is over all x_j in F such that $Q(x) \neq 0$. It is shown (L. Carlitz, *Reduction formulas for certain multiple exponential sums*, Czechoslovak Mathematical Journal, 20(95), 1970, pp. 616–627) that in general the sum $S(Q,L)$ can be expressed in terms of $K_1(a)$ or $K_2(a)$, where a is an explicit function of Q and L. More precisely, for $p = 2$ and s even, $S(Q,L)$ reduces essentially to $K_2(a)$; for s odd reduces to $K_1(a)$. For $p > 2$ and s even, $S(Q,L)$ reduces essentially to $K_2(a)$; for s odd, a variant of $K_1(a)$ is needed, namely

$$K'(a) = \sum_{\substack{u \in F \\ u \neq 0}} e(u + au^{-2}).$$

Also solved by M. G. Greening (Australia), K. S. Williams, and the proposer.

Networks with Fixed Nodes

5776 [1971, 84]. *Proposed by Jack Edmonds and Jan Mycielski, University of Colorado*

Let N be a finite network in Euclidean space, with nodes $n_1, \cdots, n_k$ and straight line (one dimensional) edges which link various pairs of nodes. Enough of the nodes are fixed in space to insure that no subset of the other nodes can be moved continuously without stretching some of the edges at the initial stages of this motion.

Prove: If some movable nodes are moved at all, there must be some edge that ends up longer than it was.

Solution by William A. Horn, National Bureau of Standards. The problem is a simple one in convexity. If n_i^1 and n_i^2, and n_j^1 and n_j^2 are, respectively, two positions for n_i and n_j, then with $1 \geqq \lambda \geqq 0$,

$$\begin{aligned}\|(\lambda n_i^1 + (1-\lambda)n_i^2) - (\lambda n_j^1 + (1-\lambda)n_j^2)\| &= \|\lambda(n_i^1 - n_j^1) + (1-\lambda)(n_i^2 - n_j^2)\| \\ &\leqq \lambda\|n_i^1 - n_j^1\| + (1-\lambda)\|n_i^2 - n_{jj}^2\|.\end{aligned}$$

Thus if $\bar{n} = (n_1, n_2, \cdots, n_k)$, and $f_{ij}(\bar{n}) = \|n_i - n_j\|$, then f_{ij} is convex. Let $\bar{n}_1$ and $\bar{n}_2$ be, respectively, the initial and final positions of the nodes, and let $\bar{m}(t) = t\bar{n}_2 + (1-t)\bar{n}_1$, $0 \leqq t \leqq 1$. Let $g_{ij}(t) = f_{ij}(\bar{m}(t))$. Then g_{ij} is also convex, and since for some pair (i,j), $g_{ij}(\tau) > g_{ij}(0)$, for τ sufficiently small it follows that

$$g_{ij}(1) \geqq g_{ij}(0) + \frac{1}{\tau}(g_{ij}(\tau) - g_{ij}(0)) > g_{ij}(0)$$

(a convex function which starts to increase is thereafter monotone). That is

$$f_{ij}(\bar{m}(1)) = f_{ij}(\bar{n}_2) > f_{ij}(\bar{m}(0)) = f_{ij}(\bar{n}_1).$$

Also solved by the proposers.

Groups Without Subgroups of Prime Index

5778 [1971, 202]. *Proposed by L. W. Shapiro, Howard University*

Find the smallest group of finite order with no subgroup of prime index.

Solution by D. M. Bloom, Brooklyn College. The alternating group A_6 is such a group. (It is simple, of order exceeding 5!, and has no proper subgroup of index $\leqq 5$.) If G is the smallest such group (excluding the trivial group $\{1\}$) and if N is a maximal normal subgroup of G, then G/N satisfies the same condition as G and hence N has order 1, G is simple. The only simple groups of composite order $\leqq 360$ are A_5 (which has a subgroup of index 5), $PSL(2,7)$ (which has a subgroup of index 7), and A_6; hence $G = A_6$.

Also solved by John Coolidge, G. A. Heuer & C. V. Heuer, J. E. Humphreys, Jim Tattersall, Z. Z. Uoiea, and the proposer.

Liouville's Theorem for Harmonic Functions

5781 [1971, 203]. *Proposed by P. R. Chernoff, University of California, Berkeley*

Generalizing the well-known result of Liouville, prove that a harmonic function $u(x)$ of polynomial growth on R^n must be a polynomial.

I. *Solution by W. C. Waterhouse, Cornell University.* Fix a point p. Let D be the mth-order partial differentiation operator $\partial^m/\partial x_{i_1} \cdots \partial x_{i_m}$, and let $M(R)$ be the maximum value of $|u|$ on the sphere of radius R about p. Writing u as a Poisson integral and differentiating under the integral, one can show

$$|Du(p)| \leqq M(R)(nm/R)^m;$$

the computation is given (for $n=3$) by O. D. Kellog, Trans. A.M.S. 33(1931), 495–496. If now $|u|$ grows no more rapidly than a polynomial of degree k, we take $m=k+1$ and let R approach infinity, concluding that $Du(p)=0$. As D and p are arbitrary, this shows u is a polynomial of degree at most k.

II. *Solution by the proposer.* Since $u(k)$ is of polynomial growth, it is a tempered distribution and its Fourier transform $\hat{u}(k)$ exists as a tempered distribution. Taking the transform of Laplace's equation $\Delta u(x)=0$, we have $|k|^2\hat{u}(k)=0$. Therefore the support of the distribution $\hat{u}(k)$ is $\{0\}$, and such distributions are finite linear combinations of derivatives of the delta function, i.e., Fourier transforms of polynomials.

Also solved by William Bosch, D. A. Hejhal, David Shelupsky, and Bertram Walsh.

Walsh shows that the hypothesis of the problem can be weakened considerably while preserving the conclusion. To conclude that a harmonic function u on R^n is a polynomial it is sufficient to assume, for example, that (1) u is of "one-sided polynomial growth," i. e., that there exist constants $K \geqq 0$ and $a \geqq 0$ such that $u(x) \leqq K\|x\|^\alpha$ when $\|x\|$ is sufficiently large (no bound on u from below); or that (2) u is of "mean polynomial growth, " i. e., for suitable K and α one has

$$\frac{1}{\sigma_n R^{n-1}} \int_{S_R} |u(x)|\, d\sigma(x) \leqq KR^\alpha$$

for all sufficiently large R.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

Hyperbolic Manifolds and Holomorphic Mappings. By Shoshichi Kobayashi. Marcel Dekker, New York, 1970. ix + 148 pp. \$11.75. (Telegraphic Review, April 1971.)

Among the interesting applications of differential geometry is its use in the study of Riemann surfaces; for example see *Introduction to Riemann Surfaces* by George Springer (1957). Kobayashi's book is a further example of this technique as applied to hyperbolic manifolds in several complex variables.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 4

CONTENTS

(*Continued on inside cover*)

APRIL 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

CERTAIN RATIONAL FUNCTIONS WHOSE POWER SERIES HAVE POSITIVE COEFFICIENTS

RICHARD ASKEY, University of Wisconsin, Madison, and
GEORGE GASPER, Northwestern University

1. Introduction. The history of mathematics is full of problems that arose in one field but had their main impact in a completely different field. One such problem was discovered in the late 1920's by K. Friedrichs and H. Lewy while working on difference approximations to the wave equation. The coefficients $A(k,m,n)$ defined by

$$(1.1)\qquad \frac{1}{(1-r)(1-s)+(1-r)(1-t)+(1-s)(1-t)} = \sum_{k,m,n=0}^{\infty} A(k,m,n)r^k s^m t^n$$

satisfy the difference equation

$$(1.2)\qquad \begin{gathered}(\Delta_k\Delta_m+\Delta_k\Delta_n+\Delta_m\Delta_n)A(k,m,n)=0,\\ k,m,n=0,1,\cdots,(k,m,n)\neq(0,0,0),\ A(0,0,0)=\tfrac{1}{3},\\ A(-1,m,n)=A(k,-1,n)=A(k,m,-1)=0,\end{gathered}$$

where $\Delta_k a(k) = a(k) - a(k-1)$.

When Δ_k is replaced by $\partial/\partial x$ the equation (1.2) takes the form

$$(1.3)\qquad \left(\frac{\partial}{\partial x}\frac{\partial}{\partial y}+\frac{\partial}{\partial x}\frac{\partial}{\partial z}+\frac{\partial}{\partial y}\frac{\partial}{\partial z}\right)u(x,y,z)=0.$$

A transformation of coordinates reduces (1.3) to the wave equation in two space dimensions. See [11] for solutions to the wave equation.

Friedrichs and Lewy hoped to use these numbers $A(k,m,n)$ to prove convergence of solutions of finite difference approximations to the wave equation to solutions of the wave equation. One fact they hoped to use was the positivity of these coefficients. Hand calculation of the first coefficients showed that they were positive. However the complete problem was surprisingly difficult. Finally Lewy wrote to G. Szegö, who was an expert on problems of this type. Szegö was able to solve the problem almost immediately [17]. The new idea Szegö had was to use the special functions of mathematical physics, in particular Bessel functions. To appreciate Szegö's idea the reader

Richard Askey received his Princeton Ph.D. in 1961 under S. Bochner. He has held positions at Washington University, the University of Chicago, and the University of Wisconsin where he is currently a Professor. He was a Guggenheim Fellow in 1969–70; his research interest is special functions.

George Gasper received his Wayne State Ph.D. in 1967 under D. Waterman, and he has held positions at the University of Wisconsin, the University of Toronto, and (currently) Northwestern University. He held a National Research Council of Canada Postdoctoral Fellowship in 1968–1969, and his research interests are Fourier analysis and special functions. *Editor.*

should spend some time trying to prove this conjecture without using special functions. It is possible [**14**], but difficult.

The proof given below uses many of Szegö's observations. It is possible, however, to replace Bessel functions by polynomials. The essential integral Szegö used was

$$\int_0^\infty J_0(ax)J_0(bx)J_0(cx)x\,dx.$$

This will be replaced by

$$\int_{-1}^1 P_k(x)P_m(x)P_n(x)\,dx,$$

where $P_n(x)$ is the Legendre polynomial. When a generalization of (1.1) is given, this integral will be replaced by

$$\cos n\theta \cos m\theta = \tfrac{1}{2}[\cos(m+n)\theta + \cos(n-m)\theta].$$

We would like to thank Prof. Szegö for calling our attention to [**17**], a paper which has fallen into undeserved obscurity, and Prof. Lewy for a helpful discussion about the background of this problem.

2. Explicit formula for the coefficients. One reason the conjecture is so hard to prove is that the variables r, s, t in (1.1) cannot be separated by factoring. A partial separation can be achieved by

$$\begin{aligned}(2.1)\quad &\frac{1}{(1-r)(1-s)+(1-r)(1-t)+(1-s)(1-t)}\\ &= \frac{1}{(1-r)(1-s)(1-t)} \cdot \frac{1}{(1/(1-r))+(1/(1-s))+(1/(1-t))}.\end{aligned}$$

The second factor in (2.1) is somewhat simpler than the left hand side and it can be changed from a sum to a product by the exponential function. Recall that

$$(2.2)\qquad \frac{1}{c} = \int_0^\infty e^{-cx}\,dx.$$

This gives

$$\begin{aligned}(2.3)\quad &\frac{1}{(1-r)(1-s)+(1-r)(1-t)+(1-s)(1-t)}\\ &= \int_0^\infty \frac{e^{-x/(1-r)}}{1-r}\,\frac{e^{-x/(1-s)}}{1-s}\,\frac{e^{-x/(1-t)}}{1-t}\,dx.\end{aligned}$$

Now r, s, and t have been factored but an integration must be performed. If $(1-r)^{-1}e^{-x/(1-r)}$ is expanded in a power series in r with coefficients power series

in x we must face the problem of trying to integrate these functions of x. Without a preliminary reduction this leads to unwieldy formulas. The simple observation

$$\frac{x}{1-r} = \frac{xr}{1-r} + x,$$

leads to

$$\frac{e^{-x/(1-r)}}{1-r} = \frac{e^{-xr/(1-r)}}{1-r} e^{-x} = e^{-x} \sum_{n=0}^{\infty} L_n(x) r^n, \tag{2.4}$$

where $L_n(x)$ is a polynomial of degree n in x. Thus (2.3) becomes

$$\frac{1}{(1-r)(1-s)+(1-r)(1-t)+(1-s)(1-t)} \tag{2.5}$$

$$= \sum \int_0^{\infty} L_k(x) L_m(x) L_n(x) e^{-3x} dx\, r^k s^m t^n,$$

so

$$A(k,m,n) = \int_0^{\infty} L_k(x) L_m(x) L_n(x) e^{-3x} dx. \tag{2.6}$$

This part of the argument is due to Szegö [**17**] as are the generalizations given in the next section.

3. Generalization to powers and more variables. Szegö observed from

$$(1-r)(1-s)+(1-r)(1-t)+(1-s)(1-t) = \frac{d}{dx}(x-r)(x-s)(x-t)\big|_{x=1}$$

that (1.1) suggests the expansion

$$\frac{1}{f'(1)} = \sum A_{n_1,\dots,n_k} x_1^{n_1} \cdots x_k^{n_k}, \tag{3.1}$$

where $f(x) = (x-x_1)\cdots(x-x_k)$. A further extension is to

$$\frac{1}{[f'(1)]^{1+\alpha}} = \sum A^{\alpha}_{n_1,\dots,n_k} x_1^{n_1} \cdots x_k^{n_k}. \tag{3.2}$$

Using a generalization of the generating function (2.4) we can obtain an explicit expression for $A^{\alpha}_{n_1,\dots,n_k}$ as an integral. Define $L_n^{\alpha}(x)$ by

$$\frac{e^{-xr/(1-r)}}{(1-r)^{\alpha+1}} = \sum_{n=0}^{\infty} L_n^{\alpha}(x) r^n, \qquad \alpha > -1. \tag{3.3}$$

A change of variables in the definition of the gamma function, $\Gamma(\alpha+1) = \int_0^{\infty} x^{\alpha} e^{-x} dx$, $\alpha > -1$, gives

$$\frac{\Gamma(\alpha+1)}{c^{\alpha+1}} = \int_0^\infty x^\alpha e^{-cx} dx, \qquad c > 0. \tag{3.4}$$

Following the argument in section 2, for $\alpha > -1$ one obtains

$$\frac{1}{[f'(1)]^{\alpha+1}} = \Sigma \frac{1}{\Gamma(\alpha+1)} \int_0^\infty L_{n_1}^\alpha(x) \cdots L_{n_k}^\alpha(x) x^\alpha e^{-kx} dx \, x_1^{n_1} \cdots x_k^{n_k}, \tag{3.5}$$

so

$$A_{n_1,\ldots,n_k}^\alpha = \frac{1}{\Gamma(\alpha+1)} \int_0^\infty L_{n_1}^\alpha(x) \cdots L_{n_k}^\alpha(x) x^\alpha e^{-kx} dx. \tag{3.6}$$

4. Linearization of the product of orthogonal polynomials. So far all that has been done is to obtain the coefficients as an integral of the product of polynomials times a weight function. This representation will not be useful until we know more about these polynomials $L_n^\alpha(x)$. A simple calculation using the generating function (3.3) gives

$$\begin{aligned}
&\sum_{m=0}^\infty \sum_{n=0}^\infty \int_0^\infty L_m^\alpha(x) L_n^\alpha(x) \, x^\alpha e^{-x} dx \, r^m s^n \\
&= \frac{1}{(1-r)^{\alpha+1}(1-s)^{\alpha+1}} \int_0^\infty x^\alpha e^{-rx/(1-r)} e^{-sx/(1-s)} e^{-x} dx \\
&= \frac{\Gamma(\alpha+1)}{\left\{(1-r)(1-s)\left[\dfrac{r}{1-r} + \dfrac{s}{1-s} + 1\right]\right\}^{\alpha+1}} \\
&= \frac{\Gamma(\alpha+1)}{(1-rs)^{\alpha+1}} = \Gamma(\alpha+1) \sum_{n=0}^\infty \binom{n+\alpha}{n} (rs)^n;
\end{aligned}$$

so that

$$\int_0^\infty L_m^\alpha(x) L_n^\alpha(x) x^\alpha e^{-x} dx = \begin{cases} 0 & , \ m \neq n \\ \dfrac{\Gamma(n+\alpha+1)}{\Gamma(n+1)}, & m = n. \end{cases} \tag{4.1}$$

The polynomials $L_n^\alpha(x)$ are called Laguerre polynomials and they have been extensively studied [18]. However, no explicit formula for the integral (3.6) is known which will allow positivity properties to be read off directly from the formula. Thus one is forced to look for an analogous integral for other orthogonal polynomials which is nonnegative and then hope to use this integral to investigate (3.6). To see what kind of integral we should be looking for, first notice when $k = 3$ and the orthogonality (4.1) is used that (3.6) gives the formal series

$$e^{-2x} L_n^\alpha(x) L_m^\alpha(x) \sim \sum_{k=0}^\infty A_{k,m,n}^\alpha \frac{\Gamma(\alpha+1)\Gamma(k+1)}{\Gamma(k+\alpha+1)} L_k^\alpha(x). \tag{4.2}$$

Later on we shall be interested in (4.2) written as

$$e^{-2x}L_n^\alpha(x)e^{-2x}L_m^\alpha(x) \sim \sum_{k=0}^{\infty} A_{k,m,n}^\alpha \frac{\Gamma(\alpha+1)\Gamma(k+1)}{\Gamma(k+\alpha+1)} e^{-2x}L_k^\alpha(x). \tag{4.3}$$

(4.3) suggests the well-known result for cos θ

$$\cos m\theta \cos n\theta = \tfrac{1}{2}\cos(n-m)\theta + \tfrac{1}{2}\cos(n+m)\theta. \tag{4.4}$$

The resemblance between (4.3) and (4.4) is heightened when we recall that $\cos n\theta = T_n(\cos\theta)$, where $T_n(x)$ is a polynomial of degree n in x, usually called the Chebychef polynomial. Also $\int_0^\pi \cos n\theta \cos m\theta \, d\theta = 0$, $m \neq n$, or $\int_{-1}^1 T_n(x)T_m(x)(1-x^2)^{-\frac{1}{2}}dx = 0$, $m \neq n$. Thus $T_n(x)$ are orthogonal polynomials and (4.4) becomes

$$T_m(x)T_n(x) = \tfrac{1}{2}T_{|n-m|}(x) + \tfrac{1}{2}T_{n+m}(x). \tag{4.5}$$

Observe that the coefficients in (4.5) are nonnegative, and by orthogonality so is

$$\int_{-1}^{1} T_k(x)T_m(x)T_n(x)(1-x^2)^{-\frac{1}{2}}dx \geqq 0. \tag{4.6}$$

There is a class of orthogonal polynomials which contains the Chebychef polynomials and also the Laguerre polynomials as limits. These are the Jacobi polynomials, $P_n^{(\alpha,\beta)}(x)$. They are orthogonal,

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)(1-x)^\alpha(1+x)^\beta dx = 0, \quad m \neq n, \quad \alpha,\beta > -1, \tag{4.7}$$

and are normalized by

$$P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n}. \tag{4.8}$$

All the facts about $P_n^{(\alpha,\beta)}(x)$ which are given without a reference are in Chapter 4 of [**18**].

$P_n^{(-\frac{1}{2},-\frac{1}{2})}(x)$ is a positive multiple of $T_n(x)$ and so (4.6) is

$$\int_{-1}^{1} P_n^{(-\frac{1}{2},-\frac{1}{2})}(x)P_m^{(-\frac{1}{2},-\frac{1}{2})}(x)P_k^{(-\frac{1}{2},-\frac{1}{2})}(x)(1-x)^{-\frac{1}{2}}(1+x)^{-\frac{1}{2}}dx \geqq 0, \ k,m,n = 0,1,\cdots.$$

In [**9**] an extension of this to $P_n^{(\alpha,\beta)}(x)$ was obtained

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(x)(1-x)^\alpha(1+x)^\beta dx \geqq 0, \tag{4.9}$$

where $\alpha \geqq \beta$, $\alpha+\beta+1 \geqq 0$, $k,m,n = 0,1,\cdots$.

The problem of finding the (α,β) for which (4.9) holds for all k, m, n was completely solved in [**10**]. All we shall need is the special case $\alpha = \beta \geqq -\frac{1}{2}$. In this case the value of the integral (4.9) as the product of Gamma functions was stated by Dougall [**5**] and a proof was given by Hsü [**13**]. The actual value is not important, but for a long time this was the only known way to prove the nonnegativity of (4.9). In addition to

the method used in [9], there is an easier method which works for $\alpha \geqq \frac{1}{2}$ (see [3]). Unfortunately this simple method does not work for $\alpha = \beta = 0$ which, after $\alpha = \beta = -\frac{1}{2}$, is probably the most important special case. The $\alpha = \beta = 0$ case was first worked out about one hundred years ago [1] and it arises in a number of different contexts. See Vilenkin [19] for some references as well as an interesting algebraic method of evaluating (4.9) in this special case. Another interesting proof is given by Dougall [6], and he also has some interesting historical comments about this result of Ferrers and Adams.

Because the $P_n^{(\alpha,\beta)}(x)$ are orthogonal, the nonnegativity of the integrals in (4.9) is equivalent to the nonnegativity of the coefficients in the expansion

$$P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x) = \sum_{k=|n-m|}^{n+m} a(k,m,n)P_k^{(\alpha,\beta)}(x) \tag{4.10}$$

(4.10) can be iterated to obtain

$$P_{n_1}^{(\alpha,\beta)}(x)\cdots P_{n_{k-1}}^{(\alpha,\beta)}(x) = \sum_{j=0}^{n_1+\cdots+n_{k-1}} a_j P_j^{(\alpha,\beta)}(x).$$

Thus if $a(k,m,n) \geqq 0$ in (4.10) for $k,m,n = 0,1,\cdots$, and some fixed (α,β), we see that

$$\int_{-1}^{1} P_{n_1}^{(\alpha,\beta)}(x)\cdots P_{n_k}^{(\alpha,\beta)}(x)(1-x)^\alpha(1+x)^\beta dx \geqq 0 \tag{4.11}$$

for the same (α,β) and $n_1,\cdots,n_k = 0,1,\cdots$.

5. Nonnegativity of the coefficients. To use (4.9) or (4.11) to investigate (3.6) there must be a way of passing from $P_n^{(\alpha,\beta)}(x)$ to $P_n^{(\gamma,\delta)}(x)$ for some (γ,δ) and also a way of going from $P_n^{(\alpha,\beta)}(x)$ to $L_n^\alpha(x)$. The passage from $P_n^{(\alpha,\beta)}(x)$ to $P_n^{(\alpha\,\beta+1)}(x)$ is particularly simple. In fact it is a general property of orthogonal polynomials.

Let $m(x)$ be a positive integrable function on $(-1,1)$ and let $p_n(x)$ be polynomials orthonormal with respect to $m(x)$, i.e.,

$$\int_{-1}^{1} p_n(x)p_m(x)m(x)dx = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$$

Also assume that $p_n(1) > 0$. This is possible since the zeros of $p_n(x)$ are all real and lie in the open integral $(-1,1)$ [18, Chapter 3]. Similarly, let $q_n(x)$ be orthonormal with respect to $(1+x)m(x)$, with $q_n(1) > 0$. Since $(1+x)q_n(x)$ is a polynomial of degree $n+1$ it may be written as

$$(1+x)q_n(x) = \sum_{k=0}^{n+1} a(k,n)p_k(x), \tag{5.1}$$

where $a(k,n) = \int_{-1}^{1}(1+x)q_n(x)p_k(x)m(x)\,dx$. But $q_n(x)$ is orthogonal to all polynomials of lower degree when integrated with respect to $(1+x)m(x)$, so $a(k,n) = 0, k = 0,1,\cdots,n-1$. Thus (5.1) becomes

$$(1 + x)q_n(x) = A_n p_{n+1}(x) + B_n p_n(x). \tag{5.2}$$

Both $p_{n+1}(x)$ and $q_n(x)$ have positive highest coefficients, since they are positive at $x = 1$ and there are no zeros to the right of $x = 1$, so $A_n > 0$. Since $p_n(x)$ has n zeros in $(-1, 1)$, $p_n(-1) = c_n(-1)^n$ with $c_n > 0$. Setting $x = -1$ in (5.2) gives

$$(-1)^{n+1}A_n c_{n+1} + B_n(-1)^n c_n = 0,$$

so $B_n = A_n c_{n+1}/c_n > 0$.

For Jacobi polynomials (5.2) becomes

$$(1 + x)P_n^{(\alpha,\beta+1)}(x) = A_n P_{n+1}^{(\alpha,\beta)}(x) + B_n P_n^{(\alpha,\beta)}(x), \ A_n > 0, \ B_n > 0. \tag{5.3}$$

Actually

$$\begin{aligned} A_n &= 2(n + 1)/(2n + \alpha + \beta + 2) \\ B_n &= 2(n + \beta + 1)/(2n + \alpha + \beta + 2), \end{aligned}$$

but these values are not important. Their positivity is all that is needed. (5.3) can be used in conjunction with (4.9) to obtain

$$\int_{-1}^{1} P_n^{(\alpha,\alpha+j)}(x) P_m^{(\alpha,\alpha+j)}(x) P_k^{(\alpha,\alpha+j)}(x)(1 - x)^\alpha(1 + x)^{\alpha+3j}dx \geqq 0, \tag{5.4}$$

$$\alpha \geqq -\tfrac{1}{2}, j = 0, 1, 2, \cdots,$$

or more generally

$$\int_{-1}^{1} P_{n_1}^{(\alpha,\alpha+j)}(x) \cdots P_{n_k}^{(\alpha,\alpha+j)}(x)(1 - x)^\alpha(1 + x)^{\alpha+kj}dx \geqq 0, \tag{5.5}$$

$$\alpha \geqq -\tfrac{1}{2}, j = 0, 1, 2, \cdots.$$

To obtain (3.6) from (5.5) we use

$$\lim_{\beta\to\infty} P_n^{(\alpha,\beta)}\left(1 - \frac{2x}{\beta}\right) = L_n^\alpha(x). \tag{5.6}$$

(5.6) can be proven from explicit expressions for $P_n^{(\alpha,\beta)}(x)$ and $L_n^\alpha(x)$

$$\begin{aligned} P_n^{(\alpha,\beta)}(x) &= \binom{n+\alpha}{n} \sum_{k=0}^{n} \frac{(-n)_k(n + \alpha + \beta + 1)_k}{(\alpha + 1)_k k!} \left(\frac{1 - x}{2}\right)^k, \\ L_n^\alpha(x) &= \binom{n+\alpha}{n} \sum_{k=0}^{n} \frac{(-n)_k}{(\alpha + 1)_k k!} x^k, \end{aligned}$$

where

$$(a)_k = a(a + 1) \cdots (a + k - 1) = \frac{\Gamma(k + a)}{\Gamma(a)}$$

or by the following general argument. When $x = 1 - 2y/\beta$, $\beta > 0$ the measure

$(1-x)^\alpha(1+x)^\beta dx$ becomes

$$\frac{-2^{\alpha+\beta+1}}{\beta^{\alpha+1}}\,y^\alpha\left(1-\frac{y}{\beta}\right)^\beta dy.$$

The factor $-2^{\alpha+\beta+1}/\beta^{\alpha+1}$ does not depend on y and so when dealing with orthonormal polynomials it can be absorbed into the polynomials. But

$$y^k y^\alpha\left(1-\frac{y}{\beta}\right)^\beta dy \to y^k y^\alpha e^{-y}dy,$$

as $\beta\to\infty$ for each $k=0,1,\cdots$, so the orthogonal polynomials for $y^\alpha(1-(y/\beta))^\beta dy$ converge to a multiple of the orthogonal polynomials for $y^\alpha e^{-y}dy$, or

$$\lim_{\beta\to\infty} P_n^{(\alpha,\beta)}\left(1-\frac{2x}{\beta}\right)=c_nL_n^\alpha(x).$$

But

$$P_n^{(\alpha,\beta)}(1)=\binom{n+\alpha}{n}=L_n^\alpha(0),$$

so $c_n=1$ and (5.6) holds.

Letting $x=1-(2y/(\alpha+j))$ in (5.4) and letting $j\to\infty$ gives

$$\int_0^\infty L_n^\alpha(x)\,L_m^\alpha(x)\,L_k^\alpha(x)x^\alpha e^{-3x}dx\geqq 0,\qquad \alpha\geqq-\tfrac{1}{2}, \tag{5.7}$$

or more generally

$$\int_0^\infty L_{n_1}^\alpha(x)\cdots L_{n_k}^\alpha(x)x^\alpha e^{-kx}dx\geqq 0,\qquad \alpha\geqq-\tfrac{1}{2}. \tag{5.8}$$

There is no problem in passing to the limit since

$$y^{j+\alpha}\left(1-\frac{y}{\beta}\right)^{3\beta}\leqq Ay^{j+\alpha}e^{-3y}$$

and the coefficients of $P_k^{(\alpha,\beta)}(1-(2y/\beta))$ converge to the coefficients of $L_k^\alpha(y)$. Since $P_kP_mP_n$ has only a finite number of terms, these two facts are all that is needed to show convergence.

The restriction $\alpha\geqq-\frac{1}{2}$ is essential, since (5.7) fails for some (k,m,n) for each $\alpha<-\frac{1}{2}$. In fact if $-1<\alpha<-\frac{1}{2}$ and r are fixed then

$$\int_0^\infty L_k^\alpha(x)\,L_m^\alpha(x)\,L_n^\alpha(x)\,x^\alpha e^{-rx}dx\geqq 0 \tag{5.9}$$

fails for some (k,m,n). The details are technical and will be given in a later paper.

6. Positivity of the coefficients and the connection with birth and death processes. Positivity can often be proven by showing that the quantity is the sum of a number

of nonnegative terms and that one of these terms is positive. Also squares are often involved. Observe that

$$\frac{1}{[f'(1)]^{2\alpha+1+1}} = \frac{1}{[f'(1)]^{\alpha+1}} \cdot \frac{1}{[f'(1)]^{\alpha+1}},$$

so

$$[\Gamma(\alpha+1)]^2 \int_0^\infty L_k^{2\alpha+1}(x)\, L_m^{2\alpha+1}(x)\, L_n^{2\alpha+1}(x)\, x^{2\alpha+1} e^{-3x} dx$$

$$= \Gamma(2\alpha+2) \sum_{a=0}^{k} \sum_{b=0}^{m} \sum_{c=0}^{n} \int_0^\infty L_{k-a}^\alpha(x)\, L_{m-b}^\alpha(x)\, L_{n-c}^\alpha(x)\, x^\alpha e^{-3x} dx \tag{6.1}$$

$$\int_0^\infty L_a^\alpha(x)\, L_b^\alpha(x)\, L_c^\alpha(x)\, x^\alpha e^{-3x} dx.$$

If $\alpha \geqq -\frac{1}{2}$ each of the terms on the right is nonnegative. Thus so is the left hand side. In particular the nonnegativity for $\alpha = -\frac{1}{2}$, which only used

$$\cos n\theta \cos m\theta = \tfrac{1}{2}[\cos(n+m)\theta + \cos(n-m)\theta], \tag{6.2}$$

implies the nonnegativity for $\alpha = 0$, so the Ferrers-Adams integral can be replaced by (6.2) as we remarked in the introduction.

A positive term can be found among the large number of terms on the right hand side of (6.1) in the following way. It is sufficient to let $a = k$, $c = 0$. Since $L_0^\alpha(x) = 1$ this term is a positive multiple of the product of two integrals of the form

$$\int_0^\infty L_n^\alpha(x)\, L_m^\alpha(x)\, x^\alpha e^{-3x} dx. \tag{6.3}$$

The positivity of (6.3) is a special case of a very important result of Karlin and McGregor [**15**].

Let $p_n(x)$ be a set of polynomials orthogonal on $[0,\infty)$ with respect to a weight function $m(x)$. Normalize $p_n(x)$ by $p_n(0) = 1$. Karlin and McGregor have shown that

$$K_\varepsilon(m,n) = \frac{\int_0^\infty p_n(x)\, p_m(x)\, e^{-\varepsilon x} m(x)\, dx}{\int_0^\infty p_n^2(x) m(x)\, dx} > 0, \qquad \varepsilon > 0 \tag{6.4}$$

$K_\varepsilon(m,n)$ has an interesting interpretation. It is the probability of moving from a population of size m to one of size n in a time interval of length ε in a birth-and-death process which is determined by the recurrence relation of $p_n(x)$. This recurrence relation has the form

$$-xp_n(x) = \alpha_n p_{n-1}(x) - (\alpha_n + \beta_n) p_n(x) + \beta_n p_{n+1}(x), \tag{6.5}$$

$\beta_n > 0$, $\alpha_{n+1} > 0$, $n = 0, 1, \cdots, \alpha_0 = 0$.

See [15] for a description of these processes as well as a proof of (6.4). Szegö outlines a proof of (6.4) in problems 81 and 82, page 386 of [18].

7. Convolution algebra. While the Friedrichs-Lewy conjecture arose in finite difference approximations to the wave equations, its most interesting application so far is to the construction of convolution algebras connected with Laguerre polynomials.

When two sequences $a(n)$, $b(m)$ are given, the usual convolution is $\sum_{-\infty}^{\infty} a(n-m)b(m)$. However, this is undefined when $a(n)$ and $b(n)$ are only defined for $n = 0, 1, \cdots$. In this case one substitute convolution is

$$\tfrac{1}{2} \sum_{m=0}^{\infty} [a(|n-m|) + a(n+m)]b(m).$$

This is connected with the product formula (4.5). Similarly it is possible to define a convolution with respect to (4.10). See [4] and [10]. Rather than repeat these definitions we shall only define a convolution with respect to (4.3). (4.3) can be written as

$$e^{-2x}\mathfrak{L}_n^\alpha(x)e^{-2x}\mathfrak{L}_m^\alpha(x) \sim \sum_{k=0}^{\infty} h^\alpha(k)B^\alpha(k,m,n)e^{-2x}\mathfrak{L}_k^\alpha(x), \tag{7.1}$$

where

$$\mathfrak{L}_k^\alpha(x) = L_k^\alpha(x)/L_k^\alpha(0),$$

$$h^\alpha(k) = \left[\int_0^\infty [\mathfrak{L}_k^\alpha(x)]^2 x^\alpha e^{-x} dx\right]^{-1},$$

and

$$B^\alpha(k,m,n) = \int_0^\infty \mathfrak{L}_k^\alpha(x)\mathfrak{L}_m^\alpha(x)\mathfrak{L}_n^\alpha(x)x^\alpha e^{-3x}dx.$$

Observe that, for $\alpha \geqq -\frac{1}{2}$,

$$\sum_{k=0}^{\infty} h^\alpha(k)|B^\alpha(k,m,n)| = \sum_{k=0}^{\infty} h^\alpha(k)B^\alpha(k,m,n) = 1. \tag{7.2}$$

The series (7.1) has only been considered as a formal series, but it actually converges for all $x \geqq 0$. However, to prove (7.2) it is only necessary to remark that the series is Abel summable at each point, [12], and so at $x = 0$

$$1 = \lim_{\rho \to 1^-} \sum_{k=0}^{\infty} \rho^k h^\alpha(k)B^\alpha(k,m,n).$$

(7.2) follows since $B^\alpha(k,m,n) \geqq 0$. Fix $\alpha \geqq -\frac{1}{2}$ and let $h(n) = h^\alpha(n)$.

For a sequence $a(n)$, $n = 0, 1, \cdots$, define

$$\| a \|_p = \left[\sum_{n=0}^{\infty} | a(n) |^p h(n) \right]^{1/p}, \qquad 1 \leqq p < \infty,$$

$$\| a \|_\infty = \sup_n | a(n) |.$$

For two sequences $a(n)$ and $b(n)$, with $\| a \|_1$ and $\| b \|_1$ dnite, define

$$c(k) = a \# b(k) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} h(m)h(n)B^\alpha(k, m, n)a(m)b(n). \tag{7.3}$$

Then

$$\begin{aligned} \| c \|_1 &= \sum h(k) | c(k) | \\ &\leqq \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} h(k)h(m)h(n) | B^\alpha(k, m, n) | \, | a(m) | \, | b(n) | \\ &= \| a \|_1 \, \| b \|_1. \end{aligned} \tag{7.4}$$

Also

$$\| c \|_\infty \leqq \| a \|_1 \, \| b \|_\infty, \tag{7.5}$$

$$\| c \|_\infty \leqq \| a \|_\infty \| b \|_1. \tag{7.6}$$

(7.4) shows that the sequences $a(n)$ with $\| a \|_1 < \infty$ form a Banach algebra under the convolution (7.3). (7.5) and (7.6) added to (7.4) give a convolution algebra. Among other results we have Young's inequality

$$\| c \|_r \leqq \| a \|_p \| b \|_q, \quad \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1, \qquad 1 \leqq p, q, r \leqq \infty.$$

See [16] for an axiomatic treatment of convolution algebras.

Associate to the sequence $a(n)$ with $\| a \|_1$ finite its Laguerre series in the form

$$f(x) = \sum_{n=0}^{\infty} a(n) h(n) \mathfrak{L}_n^\alpha(x) e^{-2x}. \tag{7.7}$$

If $g(x) = \sum_{n=0}^{\infty} b(n) h(n) \mathfrak{L}_n^\alpha(x) e^{-2x}$, $\| b \|_1 < \infty$, and $c(k) = a \# b(k)$ is defined by (7.3), then $\| c \|_1 < \infty$ and

$$f(x) g(x) = \sum_{k=0}^{\infty} c(k) h(k) \mathfrak{L}_k^\alpha(x) e^{-2x} \tag{7.8}$$

or the Laguerre series of the convolution of two sequences is the product of their Laguerre series. This operational property is the reason the convolution (7.3) is defined this way. There are many other ways to define a convolution so that a property like (7.8) holds. For example, if instead of (7.7) the Laguerre series are

$$f(x) \sim \sum_{n=0}^{\infty} a(n)h(n)\mathfrak{L}_n^\alpha(x),$$

$$g(x) \sim \sum_{n=0}^{\infty} b(n)h(n)\mathfrak{L}_n^\alpha(x),$$

then

$$f(x)\,g(x) \sim \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a(n)\,b(m)\,h(n)\,h(m)\,\mathfrak{L}_n^\alpha(x)\,\mathfrak{L}_m^\alpha(x)$$

$$= \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} h(k)\,h(m)\,h(n)\,a(n)\,b(m)\,C^\alpha(k,m,n)\,\mathfrak{L}_k^\alpha(x),$$

where $\mathfrak{L}_n^\alpha(x)\,\mathfrak{L}_m^\alpha(x) = \sum_{k=0}^{n+m} C^\alpha(k,m,n)\,h(k)\,\mathfrak{L}_k^\alpha(x)$. Thus

$$c(k) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} C^\alpha(k,m,n)\,a(m)\,b(n)\,h(m)\,h(n)$$

is a convolution. However, the above proof of an inequality like (7.4) fails. $\sum_{k=0}^{\infty} C^\alpha(k,m,n)h(k) = 1$ still holds, but it is no longer true that $\sum_{k=0}^{\infty} |C^\alpha(k,m,n)|h(k) = 1$. In fact $(-1)^{k+m+n}C^\alpha(k,m,n) \geqq 0$, $\alpha > -1$, (see [3], [7]).

In addition to the norm inequalities it is possible to ask many of the questions of probability theory in this context. Given two distribution sequences $a(n)$, $b(n)$, i.e., $a(n) \geqq 0$, $\sum_{n=0}^{\infty} a(n) = 1$, the sequence $c(n)$ defined by (7.3) is also a distribution sequence. Probability theory questions can be asked in this context. For example, what are the infinitely divisible and stable laws and what is the central limit theorem? Various harmonic analysis questions can also be asked. For example, is the converse of the Wiener-Lévy theorem true, as it is for the algebra of absolutely convergent Fourier series with the usual convolution?

8. Further results and open problems. As was pointed out in section 4, the linearization results for Jacobi polynomials can be iterated to obtain new positivity results with more variables. In particular (4.9) implies (4.11). A similar result is also true for Laguerre polynomials but something is lost in the process.

$$\int_0^\infty L_n^\alpha(x)\,L_m^\alpha(x)\,L_k^\alpha(x)\,x^\alpha e^{-3x}dx \geqq 0, \qquad \alpha \geqq -\tfrac{1}{2} \tag{8.1}$$

is equivalent to

$$e^{-2x}L_n^\alpha(x)\,e^{-2x}L_m^\alpha(x) = \sum_{k=0}^{\infty} D^\alpha(k,m,n)\,e^{-2x}L_k^\alpha(x),\ D^\alpha(k,m,n) \geqq 0, \tag{8.2}$$

and (8.2) can be iterated to obtain

$$e^{-2x}L_j^\alpha(x)\,e^{-2x}L_n^\alpha(x)\,e^{-2x}L_m^\alpha(x) = \sum_{k=0}^{\infty} D(k,j,m,n)\,e^{-2x}L_k^\alpha(x), \qquad D(k,j,m,n) \geqq 0. \tag{8.3}$$

(8.3) is equivalent to

$$\int_0^\infty L_n^\alpha(x)\,L_m^\alpha(x)\,L_k^\alpha(x)\,L_j^\alpha(x)\,x^\alpha e^{-5x}dx \geqq 0, \qquad \alpha \geqq -\tfrac{1}{2}. \tag{8.4}$$

However, (8.4) is a weaker result than (5.8) for $k = 4$, i.e.,

$$\int_0^\infty L_n^\alpha(x)\,L_m^\alpha(x)\,L_k^\alpha(x)\,L_j^\alpha(x)\,x^\alpha e^{-4x}dx \geqq 0, \qquad \alpha \geqq -\tfrac{1}{2}. \tag{8.5}$$

The Karlin-McGregor result (6.4) can be used to show that (8.5) implies (8.4), but infinite series are needed and there is another way which only uses finite series. Letting $4x = 5y$ gives for (8.5)

$$\int_0^\infty L_n^\alpha\left(\frac{5y}{4}\right)L_m^\alpha\left(\frac{5y}{4}\right)L_k^\alpha\left(\frac{5y}{4}\right)L_j^\alpha\left(\frac{5y}{4}\right)y^\alpha e^{-5y}dy \geqq 0.$$

Then using $(\lambda)^n\mathfrak{L}_n^\alpha(\lambda^{-1}x) = \sum_{k=0}^n\binom{n}{k}\,(\lambda-1)^{n-k}\mathfrak{L}_k^\alpha(x)$ [**18**, problem 67] for $\lambda = 5/4$, $x = 5y/4$ gives (8.4). This suggests there should be a stronger result than (8.1). Such a result would follow from $e^{-x}L_n^\alpha(x)e^{-x}L_m^\alpha(x) = \sum_{k=0}^\infty E(k,m,n)e^{-x}L_k^\alpha(x)$, $E(k,m,n) \geqq 0$. This is equivalent to

$$\int_0^\infty L_n^\alpha(x)\,L_m^\alpha(x)\,L_k^\alpha(x)\,x^\alpha e^{-2x}dx \geqq 0 \tag{8.6}$$

and it can be iterated to obtain

$$\int_0^\infty L_{n_1}^\alpha(x)\,L_{n_2}^\alpha(x)\cdots L_{n_k}^\alpha(x)\,x^\alpha e^{-(k-1)x}dx \geqq 0. \tag{8.7}$$

In another paper we shall show that (8.6) holds for $k, m, n = 0, 1, \cdots$ when $\alpha \geqq (-5+\sqrt{17})/2$ and the only case of equality is

$$\int_0^\infty [L_1^{\alpha_0}(x)]^3 x^{\alpha_0}e^{-2x}dx = 0, \qquad \alpha_0 = (-5+\sqrt{17})/2.$$

Observe that α_0 is slightly larger than $-\frac{1}{2}$. A calculation of the type given in section 2 shows that (8.6) is equivalent to the nonnegativity of the coefficients in

$$\begin{aligned}&\frac{1}{[(1-r)(1-s)(2+t)+(1-r)(2+s)(1-t)+(2+r)(1-s)(1-t)]^{\alpha+1}}\\ &\quad= \sum E^\alpha(k,m,n)r^\lambda s^m t^n.\end{aligned} \tag{8.8}$$

This time the series (8.8) is examined, a different representation for $E^\alpha(k,m,n)$ is found as a sum and the positivity of this sum is proven using transformation formulas and recurrence relations for generalized hypergeometric series. Again special functions are crucial, but this time a more complicated set of functions, generalized hypergeometric functions ${}_3F_2(a,b,c;\,d,e;\,1)$.

H. Lewy suggested to one of us that $1/a$, where

$$a = (4 - r - s - t - u)[(1 - r)(1 - s) + (1 - r)(1 - t) + (1 - r)(1 - u) + (1 - s)(1 - t) + (1 - s)(1 - u) + (1 - t)(1 - u)] \tag{8.9}$$

should have positive power series coefficients. Without the factor $1/(4 - r - s - t - u)$ these coefficients should approach zero, since the coefficients satisfy finite difference approximations to the wave equation in three space dimensions and Huygen's principle holds in three space. The factor $1/(4 - r - s - t - u)$ is an averaging factor which counts the earlier terms more than the later terms. It is easy to show that the early terms are positive so that an averaging should give positive coefficients. Unfortunately we have been unable to obtain a representation for the coefficients in (8.9) in a useful form. This suggests that there should be other ways of proving the positivity of the coefficients in (1.1) and (8.8) which will extend to (8.9). There is one other way of proving positivity due to Kaluza [**14**]. However, his proof only works for three variables and $\alpha = 0$, i.e., only for (1.1), and even in this case his proof is quite difficult. He does obtain some monotonicity results which do not follow from the above arguments. There should be a combinatorial interpretation of these results and if so this might suggest new methods.

References

1. J. C. Adams, Proc. Roy. Soc., 27, (1878) 63; also Collected Scientific Papers I, p. 187.

2. R. Askey, Orthogonal polynomials and positivity, in Studies in Applied Mathematics, Wave Propagation and Special Functions, SIAM, 1970, 64–85.

3. ———, Linearization of the product of orthogonal polynomials, *Studies in Analysis*, papers dedicated to S. Bochner, edited by R. Gunning, Princeton Univ. Press, Princeton, 1971, pp. 131–138.

4. R. Askey, and S. Wainger, A dual convolution structure for Jacobi polynomials, *Orthogonal Expansions and their Continuous Analogues*, edited by D. Haimo, Southern Illinois Univ. Press, Carbondale and Edwardsville, Illinois, 1968, 25–36.

5. J. Dougall, A theory of Sonine in Bessel functions, with two extensions to spherical harmonics, Proc. Edinburgh Math. Soc., 37 (1919) 33–47.

6. ———, The product of two Legendre polynomials, Proc. Glasgow Math. Assoc., 1 (1952-55) 121–125.

7. A. Erdélyi, On some expansions in Laguerre polynomials, Jour. London Math. Soc., 13 (1938) 154–156.

8. ———, Higher Transcendental Functions, Vol. 2, McGraw-Hill, New York, 1953.

9. G. Gasper, Linearization of the product of Jacobi polynomials, I, Canad. J. Math., 22 (1970) 171–175.

10. ———. Linearization of the product of Jacobi polynomials, II, Canad. J. Math., 22 (1970) 582–593.

11. J. Hadamard, Lectures on Cauchy's Problem in Linear Differential Equations, Dover, New York, 1952.

12. E. Hille, On Laguerre series, second note, Proc. Nat. Acad. Sci. (Washington, D. C.) 12 (1926) 265–269.

13. H. -Y. Hsü, Certain integrals and infinite series involving ultraspherical polynomials and Bessel functions, Duke Math. J., 4 (1938) 374–383.

14. Th. Kaluza, Elementarer Beweis einer Vermutung von K. Friedrichs und H. Lewy, Math. Z., 37 (1933) 689–697.

15. S. Karlin and J. McGregor, The differential equations of birth-and-death processes, and the Stieltjes moment problem, Trans. Amer. Math. Soc., 85 (1957) 489–546.

16. R. O'Neil, Convolution operators and $L(p, q)$ spaces, Duke Math. J., 30 (1963) 129–142.

17. G. Szegö, Über gewisse Potenzreihen mit lauter positiven Koeffizienten, Math. Z., 37 (1933) 674–688.

18. ———, Orthogonal Polynomials, Colloquium Publications, AMS, 23 (1959).

19. N. Ya. Vilenkin, Special Functions and the Theory of Group Representations, Translations of Mathematical Monographs, AMS, 22 (1968).

REMARKS ON THE LEBESGUE DIFFERENTIATION THEOREM, THE VITALI LEMMA AND THE LEBESGUE-RADON-NIKODYM THEOREM

MIGUEL DE GUZMÁN AND BALDEMERO RUBIO, Universidad de Madrid.

Introduction. The purpose of this paper is to give simple proofs of certain theorems of elementary measure theory. The proofs we offer present certain elements of novelty and admit rather trivial extensions to more general situations. However, in order to make our exposition as simple as possible, we place ourselves in an elementary context, where the theorems and proofs can be presented in an easy way.

In Section 1 we present two easy covering lemmas, which are then used in Section 2 to obtain a simple proof of the Lebesgue differentiation theorem and in Section 3 to give a version of the Vitali lemma that is valid for rather general measures. In Section 4 we present a simple proof of the Lebesgue-Radon-Nikodym theorem.

We wish to thank G. Weiss for his advice concerning the arrangement of this paper. The work of the first named author was supported by the Fundación Juan March, Madrid, and the Institut Mittag-Leffler, Djursholm, Sweden.

1. Two covering lemmas. The two following simple lemmas are of a purely geometrical character. The first one can replace the Vitali lemma, as shown in Section 2, in the proof of the Lebesgue differentiation theorem, providing a very easy approach to it. The second one gives a simple proof of the Vitali lemma for general measures, as will be shown in Section 3.

1.1. LEMMA. *Let $\{R_i\}_1^k$ be a finite sequence of closed intervals of $\boldsymbol{R}^n$, centered at the origin and with non-empty interiors. Assume also $R_1 \subseteq R_2 \subseteq \cdots \subseteq R_k$. Let A be a bounded set of $\boldsymbol{R}^n$. For each $x \in A$, select an integer $i(x)$, $1 \leqq i(x) \leqq k$, and write $R(x) = x + R_{i(x)}$. Then there exists a finite subset $x_1, x_2, \cdots, x_h$ of elements*

of A such that $A \subseteq \bigcup_{j=1}^{h} R(x_j)$ and each $y \in \boldsymbol{R}^n$ is at most in 2^n of these sets $R(x_j)$.

Proof. Choose x_1 such that $i(x_i)$ is as large as possible. Assume $x_1, x_2, \cdots, x_m$ have been already chosen. Then take $x_{m+1} \in A - \bigcup_{j=1}^{m} R(x_j)$ such that $i(x_{m+1})$ is as large as possible. Since A is bounded and since the sets $R(x_j)$ we thus obtain are such that

$$R^*(x_j) = x_j + \tfrac{1}{2} R_{i(x_j)}$$

are obviously disjoint, we end this selection process in a finite number d of steps, obtaining $A \subseteq \bigcup_{j=1}^{h} R(x_j)$. We now prove that any $y \in \boldsymbol{R}^n$ is at most in 2^n sets $R(x_j)$. To see this, draw n hyperplanes through y parallel to the coordinate hyperplanes and consider the 2^n closed quadrants around y so obtained. In each quadrant there is at most one x_j with $y \in R(x_j)$. For if there were two, the larger $R(x_j)$ would contain the center of the smaller one, and this is excluded by construction. This proves the lemma.

1.2. LEMMA. *Let $\{R_k\}$ be a sequence of closed intervals of $\boldsymbol{R}^n$, centered at the origin and with non-empty interiors. Assume also $R_1 \supseteq R_2 \supseteq \cdots$ and $\bigcap R_k = \{0\}$. Let A be a bounded set of $\boldsymbol{R}^n$. For each $x \in A$, take a positive integer $i(x)$ and write $R(x) = x + R_{i(x)}$. Then there exists a sequence $\{x_k\} \subseteq A$ such that*

(a) $A \subseteq \bigcup R(x_k)$,

(b) *each $y \in \boldsymbol{R}^n$ is in at most 2^n of the sets $R(x_k)$,*

(c) *the sets $R(x_k)$ can be distributed into $4^n + 1$ disjoint sequences.*

Proof. Take x_1 such that $R(x_1)$ is as big as possible. Assume $x_1, x_2, \cdots, x_m$ have been already chosen. If $A - \bigcup_{k=1}^{m} R(x_k) = \varnothing$ the process of selection stops. Otherwise we take $x_{m+1} \in A - \sum_{k=1}^{m} R(x_k)$ such that $R(x_{m+1})$ is as big as possible. The sequence obtained in this way satisfies the following properties: (1) If $i \neq j$ then the center of $R(x_i)$ is outside $R(x_j)$; (2) the sequence of numbers {side length $R(x_k)$} either is finite or its limit is zero as $k \to \infty$, since the sets $x_k + \frac{1}{2} R_{i(x_k)}$ are disjoint. If the selection process stops, then (a) is trivial. Otherwise, if $x \in A - \bigcup R(x_k)$, there exists a j such that the side length of $R(x)$ is greater than that of $R(x_j)$, and this means that $R(x)$ has been overlooked in our selection.

The proof of (b) is the same as in Lemma 1.1.

In order to prove (c), fix an element $R(x_j)$ of the sequence $\{R(x_k)\}$. According to (b), no more than 2^n elements of $\{R(x_k)\}$ contain a fixed vertex of $R(x_j)$. Now any $R(x_k)$ with $k < j$ is not smaller in size than $R(x_j)$, so if

$$R(x_k) \cap R(x_j) \neq \varnothing, \qquad k < j,$$

then $R(x_k)$ contains at least one vertex of $R(x_j)$. Hence for each j, no more than 4^n elements of the set $\{R(x_1), \cdots, R(x_{j-1})\}$ can have non-empty intersections with $R(x_j)$. This fact permits us to distribute the sets $R(x_k)$ into $4^n + 1$ disjoint sequences

$I_1, I_2, \cdots, I_{4^n+1}$ in the following manner: Take $R(x_i) \in I_i$ for $i = 1, 2, \cdots, 4^n + 1$. Since $R(x_{4^n+2})$ is disjoint with $R(x_k)$ for some $k \leqq 4^n + 1$, we can set $R(x_{4^n+2}) \in I_k$, etc. This proves (c).

2. The Lebesgue differentiation theorem. Lemma 1.1, together with the continuity of the integral and the Heine-Borel theorem, yield the Lebesgue differentiation theorem in the following general form:

2.1. THEOREM. *Let $f \in L^1(\boldsymbol{R}^n)$. Consider any collection $\mathscr{R}$ of closed intervals with non-empty interior, centered at the origin 0 and containing sequences $\{R_k\}$ contracting to 0 as $k \to \infty$. (This will be denoted $R_k \to 0$ as $k \to \infty$.) Assume furthermore that the intervals in $\mathscr{R}$ are comparable, i.e., for any two $R_1, R_2 \in \mathscr{R}$ either $R_1 \subseteq R_2$ or $R_2 \subseteq R_1$. (Example: the collection of all closed cubic intervals centered at 0.)*

For $x \in \boldsymbol{R}^n$ and $\{R_k\} \subseteq \mathscr{R}$ write $R_k(x) = x + R_k$. Then, for almost every $x \in \boldsymbol{R}^n$ and for every $\{R_k\} \subseteq \mathscr{R}$ with $R_k \to 0$ as $k \to \infty$, one has

$$\lim_{k \to \infty} \frac{1}{|R_k(x)|} \int_{R_k(x)} f(y)dy = f(x).$$

Proof. Define

$$Mf(x) = \sup_{R \in \mathscr{R}} \frac{1}{|R|} \int_{R(x)} |f(y)| \, dy.$$

The function Mf is measurable, since $\{x : Mf(x) > \lambda\}$ is an open set. The operator M is called the **Hardy-Littlewood maximal operator** associated with $\mathscr{R}$. For any $f \in L^1(\boldsymbol{R}^n)$ and any $\alpha > 0$, we shall show that

$$|\{x : Mf(x) > \alpha\}| \leqq \frac{2^n}{\alpha} \|f\|_1.$$

In fact, consider any compact subset K of

$$\{x : Mf(x) > \alpha\}.$$

If $x \in K$ there is an interval $R \in \mathscr{R}$ such that $\int_{R(x)} |f(y)| \, dy > \alpha |R|$.

By the continuity of the integral, there is a neighborhood $U(x)$ of x so that if $z \in U(x)$, then

$$\frac{1}{|R|} \int_{R(z)} |f(y)| \, dy > \alpha$$

for the same R as before.

Since K is compact, we can choose a finite number of such neighborhoods $U(x)$ covering K. Hence there is a finite set $\{R_i\}_1^k$ of intervals of $\mathscr{R}$ such that for each $x \in K$ we can choose an index $i(x) \in \{1, 2, \cdots, k\}$ in such a way that, if $S(x) = x + R_{i(x)}$,

then $\int_{S(x)} |f(y)| dy > \alpha |S(x)|$. We now apply Lemma 1.1 and choose $\{S(x_j)\}$. If χ_M denotes the characteristic function of the set M, then we clearly have

$$|K| \leqq \Sigma |S(x_j)| \leqq \frac{1}{\alpha} \Sigma \int_{S(x_j)} |f(y)| dy$$

$$= \frac{1}{\alpha} \int |f(y)| \Sigma \chi_{S(x_j)}(y) dy \leqq \frac{2^n}{\alpha} \| f \|_1,$$

independent of K. This proves that

$$|\{x : Mf(x) > \alpha\}| \leqq (2^n/\alpha) \| f \|_1 .$$

The theorem is now an easy consequence of this fact. Define

$$\bar{D}(\int f, \mathscr{R}, x) = \bar{D}(\int f, x) = \sup \overline{\lim_{k \to \infty}} \frac{1}{|R_k|} \int_{R_k(x)} f(y) dy,$$

where the sup is taken over all sequences $\{R_k\} \subseteq \mathscr{R}$ with $R_k \to 0$ as $k \to \infty$. We define $\underline{D}(\int f, \mathscr{R}, x) = \underline{D}(\int f, x)$ similarly using inf lim. We shall prove that for fixed $\alpha > 0$, we have $|E_\alpha| = |\{x : |\bar{D}(\int f, x) - f(x)| > \alpha\}| = 0$. In fact, take any $\varepsilon > 0$ and set $f = g + h$, where g is a continuous function and h is in $L^1(\mathbf{R}^n)$ with $\| h \|_1 \leqq \varepsilon$. Then

$$E_\alpha = \{x : |\bar{D}(\int h, x) - h(x)| > \alpha\} \subseteq \left\{x : |\bar{D}(\int h, x)| > \frac{\alpha}{2}\right\} \cup \left\{x : |h(x)| > \frac{\alpha}{2}\right\}.$$

Call A_1 and A_2 the two sets in the last member of the preceding relationship. Now $A_1 \subseteq \{x : Mh(x) > \frac{1}{2}\alpha\}$ and $|\{x : Mh(x) > \frac{1}{2}\alpha\}| \leqq (2^{n+1}/\alpha)\varepsilon$. As for A_2, we have $|A_2| = \int_{A_2} dy \leqq (2/\alpha) \| h \|_1 \leqq (2/\alpha)\varepsilon$. Since ε is arbitrarily small, it is easy to see that $|E_\alpha| = 0$. In the same way, $|\{x : |\underline{D}(\int f, x) - f(x)| > \alpha\}| = 0$. This relation easily yields

$$|\{x : \bar{D}(\int f, x) \neq f(x) \text{ or } \underline{D}(\int f, x) \neq f(x)\}| = 0,$$

and this proves the theorem.

3. A general form of the Vitali lemma. The use of Lemma 2.2 in order to obtain the Vitali lemma for a rather general measure is mainly based on property (c) of the selected covering in that lemma.

3.1. THEOREM. *Let $\{R_k\}$ be a sequence of non-increasing closed intervals of $\mathbf{R}^n$, with non-empty interiors, centered at the origin 0, and contracting to 0 as $k \to \infty$. Let μ be a non-negative measure defined on the Lebesgue measurable sets of $\mathbf{R}^n$. Let P be any bounded measurable set with $\mu(P) < \infty$. For each $x \in P$, let $\{S_j(x)\}$ be a sequence of translations of elements of $\{R_k\}$ centered at x with $S_j(x) \to x$ as $j \to \infty$. Then we can select from $\{S_j(x) : x \in P,\ j = 1, 2, \cdots\}$ a sequence of disjoint sets $\{T_k\}$ such that $\mu(P - \bigcup T_k) = 0$.*

Proof. For each $x \in P$ take an arbitrary element of the sequence $\{S_j(x)\}$. Apply now Lemma 1.2 to obtain a sequence $\{A_k\}$ of intervals covering P which can be distributed into $4^n + 1$ disjoint sequences. At least for one of these sequences, call it $\{B_k\}$, we have $\mu[(\bigcup B_k) \cap P] \geqq (4^n + 1)^{-1}\mu(P)$. Otherwise $\mu[(\bigcup A_k) \cap P] < \mu(P)$ so $\bigcup A_k$ does not cover P. Taking a finite number of elements of $\{B_k\}$, call them $\{T_k\}_1^{h_1}$, we can still have

$$\mu\left[\left(\bigcup_{k=1}^{h_1} T_k\right) \cap P\right] \geqq \frac{1}{4^n + 2}\mu(P).$$

Call $a = 1 - 1/(4^n + 2)$. We have $0 < a < 1$, $\mu(P - \bigcup_{k=1}^{h_1} T_k) \leqq a\mu(P)$. We repeat the process with $P_1 = P - \bigcup_{k=1}^{h_1} T_k$, but taking now from each sequence $\{S_j(x)\}$ for $x \in P_1$ an interval that is disjoint from $\bigcup_{k=1}^{h_1} T_k$. We get now $\{T_k\}_{h_1+1}^{h_2}$ and

$$\mu\left(P - \bigcup_{k=1}^{h_2} T_k\right) \leqq a^2\mu(P).$$

This whole process can be finished in finite number of steps or can be infinite, but in any case we obtain $\mu(P - \bigcup T_k) = 0$.

4. The Lebesgue-Radon-Nikodym thèorem. We shall consider subsets of the open unit cube $Q_0 = \{x \in \mathbf{R}^n : 0 < x_i < 1, i = 1,2,\cdots,n\}$ in $\mathbf{R}^n$. In Q_0 we define the system D of open "dyadic cubes." An open dyadic cube of side length 2^{-k}, $k = 0,1,2,\cdots$, is a set of the form

$$Q_k = \{x \in Q_0 : h_i 2^{-k} < x_i < (h_i + 1)2^{-k},\ i = 1,2,\cdots,n\},$$

where $h_i \in N$, $0 \leqq h_i \leqq 2^k - 1$. Denote by Q_0' the set of points of Q_0 which do not belong to the boundary of any dyadic cube. We clearly have $|Q_0'| = |Q_0|$. Obviously, given any two dyadic cubes, either they are disjoint or one is contained in the other. This property makes the proof of the following covering lemma easy.

4.1. LEMMA. *Let A be a subset of Q_0'. Assume that for each $x \in A$ we are given a dyadic cube $Q(x)$ containing x. Then we can choose a disjoint sequence $\{T_k\}$ of such cubes so that $A \subseteq \bigcup T_k$.*

It is enough to make the selection by the order of the side lengths of the given cubes, beginning with the big ones and excluding at each step the cubes which have been already covered.

This lemma leads us to a natural way of proving the following form of the Lebesgue-Radon-Nikodym theorem:

4.2. THEOREM. *Let μ be a nonnegative finite measure defined on the Lebesgue measurable subsets of the unit cube Q_0. Assume that μ is absolutely continuous with respect to λ, the Lebesgue measure. Then there is a function $g \in L^1(Q_0)$ such that $\mu(P) = \int_P g(y)dy$ for each Lebesgue set $P \subseteq Q_0$.*

Proof. First we use the Hahn decomposition theorem to obtain a disjoint sequence of measurable sets $E_0, E_1, E_2, \cdots$, such that $Q_0 = \bigcup E_j$, $|E_0| = 0$ and $(j-1)|M| \leqq \mu(M) \leqq j|M|$ for each measurable set $M \subseteq E_j$ with $j \geqq 1$. For this purpose, we consider first the measure $\lambda - \mu$. By the Hahn theorem there exists a measurable set E_1 such that $\mu(M) \leqq |M|$ for every measurable $M \subseteq E_1$ and $\mu(M) \geqq |M|$ for every $M \subseteq Q_0 - E_1$. Consider the measure $2\lambda - \mu$ in $Q_0 - E_1$. There is a measurable set $E_2 \subseteq Q_0 - E_1$ such that

$$|M| \leqq \mu(M) \leqq 2|M|$$

for each measurable set $M \subseteq E_2$ and $\mu(M) \geqq 2|M|$ for $M \subseteq Q_0 - (E_1 \cup E_2)$, etc. Let $E_0 = Q_0 - \bigcup_{j \geqq 1} E_j$. Then for each $j \geqq 1$, we have $\infty \geqq \mu(Q_0) \geqq \mu(E_0) \geqq j|E_0|$. Hence $|E_0| = 0$. Now for fixed j and k, consider the following function:

$$f_{jk}(x) = \sum_i \frac{\mu(Q_k^i \cap E_j)}{|Q_k^i \cap E_j|} \chi_D(x), \quad D = Q_k^i \cap E_j,$$

where the sets Q_k^i are the dyadic cubes with side length 2^{-k} and χ_D denotes the characteristic function of the set D. The fractions in the definition of f_{jk} are taken to be zero in case the denominator (and consequently also the numerator) is zero. Observe that $f_{0k}(x) = 0$ almost everywhere for all k. This will permit us to restrict the following argument to $j \geqq 1$.

We next prove for every fixed $j \geqq 1$, that

$$\lim_{k \to \infty} f_{jk}(x) = g_j(x)$$

exists almost everywhere in Q_0. It is obvious from the definition of E_j and of f_{jk} that for every $x \in Q_0$ we have $f_{jk} \leqq j$ for j fixed and k arbitrary. Once we have established this, we have, for every finite union A of disjoint dyadic cubes,

$$\int_{A \cap E} f_{jk}(y)dy = \mu(A \cap E_j)$$

(k sufficiently large). Hence, by the bounded convergence theorem, for every j,

$$\int_{A \cap E_j} g_j(y)dy = \mu(A \cap E_j).$$

Also if we call $g = \sum g_j$, then $g \in L^1(Q_0)$ and

$$\int_A g(y)dy = \mu(A).$$

For an arbitrary measurable $P \subseteq Q_0$, we take $\{A_k\}$ each A_k being a finite union of disjoint dyadic cubes, such that $|P - A_k| + |A_k - P| \to 0$ as $k \to \infty$. Then, by the continuity of μ and of the integral,

$$\mu(P) = \lim_{k\to\infty} \mu(A_k) = \lim_{k\to\infty} \int_{A_k} g(y)dy = \int_P g(y)dy.$$

This proves the theorem.

Thus it only remains to prove for every fixed $j \geqq 1$, that $\lim_{k\to\infty} f_{jk}(x)$ exists almost everywhere in Q_0. Observe first that for each $x \notin E_j$ we have $f_{jk}(x) = 0$ for all k, and that

$$\lim_{k\to\infty} f_{jk}(x) = \lim_{k\to\infty} \frac{\mu(Q_k(x) \cap E_j)}{|Q_k(x) \cap E_j|}$$

for each $x \in E_j \cap Q_0'$, where the sets $Q_k(x)$ are the dyadic cubes containing x and contracting to x as $k \to \infty$. We shall call $E_j = B$ and we wish to prove

$$\bar{D}(\mu, x) \equiv \overline{\lim_{k\to\infty}} \frac{\mu(Q_k(x) \cap B)}{|Q_k(x) \cap B|} = \underline{\lim}_{k\to\infty} \frac{\mu(Q_k(x) \cap B)}{|Q_k(x) \cap B|} \equiv \underline{D}(\mu, x)$$

for almost all $x \in B \cap Q_0'$. It is clear that $\bar{D}(\mu, \cdot)$ is measurable and $\leqq j$. Consider the set

$$C_{rs} = \{x \in Q_0' \cap B : \bar{D}(\mu, x) > r > s > \underline{D}(\mu, x)\}$$

for rational numbers $r > s > 0$. We shall prove $|C_{rs}| = 0$ and this gives the result we seek.

In fact, take any open set $G \supset C_{rs}$. By definition of C_{rs} and by Lemma 4.1. it is clear that we can select two sequences $\{T_k\}$ and $\{T_k'\}$ of disjoint dyadic cubes satisfying: $T_k \subseteq G$, $T_k' \subseteq G$, $\mu(T_k \cap B) > r|T_k \cap B|$, $\mu(T_k' \cap B) < s|T_k' \cap B|$, $C_{rs} \subseteq \bigcup (T_k \cap B)$, and $C_{rs} \subseteq \bigcup (T_k' \cap B)$. Therefore we have

$$\mu(G) \geqq \mu(\bigcup T_k) \geqq \sum \mu(T_k \cap B) > \sum r|T_k \cap B| \geqq r|C_{rs}|$$

and

$$\mu(C_{rs}) \leqq \mu[\bigcup (T_k' \cap B)] \leqq \sum \mu(T_k' \cap B) < s|\bigcup (T_k' \cap B)| \leqq s|C_{rs}|.$$

Since μ is absolutely continuous, given $\varepsilon > 0$ we can choose G so that $|G| - |C_{rs}| \leqq \varepsilon$ and $\mu(G) - \mu(C_{rs}) \leqq \varepsilon$. Thus we obtain

$$|C_{rs}| \leqq \frac{1}{r}\mu(G) \leqq \frac{1}{r}(\varepsilon + \mu(C_{rs})) \leqq \frac{1}{r}(\varepsilon + s|G|) \leqq \frac{1}{r}[\varepsilon + s(\varepsilon + |C_{rs}|)].$$

Hence $|C_{rs}| \leqq \varepsilon(1+s)/(r-s)$. Since ε is arbitrarily small, $|C_{rs}| = 0$.

5. Remarks. We have not been able to find in the literature the use of Lemma 1.1, in combination with the Heine-Borel theorem to prove the Lebesgue differentiation theorem. The lemma has been used in [4] as a substitute for a well-known lemma of Calderón and Zygmund [2] in the theory of singular integral operators. Cotlar [3] has introduced the use of this type of "almost disjoint" covering lemma for cubes in the classical theory of the Hardy-Littlewood maximal operator and in singular integrals.

The idea of the proof of the Vitali lemma we present in Section 3 has its origin in Besicovitch [1], who was the first in considering this kind of lemma for spheres in connection with the Vitali lemma and its generalizations.

References

1. A. S. Besicovitch, A general form of the covering principle and relative differentiation of additive functions, Proc. Cambridge Phil. Soc., 41 (1945) 103–110.

2. A. P. Calderón and A. Zygmund, On the existence of certain singular integrals, Acta Math., 88 (1952) 85–139.

3. M. Cotlar, A general interpolation theorem for linear operations, Revista Matemática Cuana, 1 (1955) 57–84.

4. M. de Guzmán, A covering lemma with applications to differentiability of measures and singular integral operators, Studia Math., 34 (1970) 299–317.

THE NONLINEAR SIMPLE PENDULUM

FRED BRAUER, University of Wisconsin

1. It is customary in elementary courses in differential equations to derive a mathematical model for the motion of a simple pendulum released from rest at a given angle with the vertical (see, for example [1, Section 1.2]). If it is assumed that the only external forces are a constant gravitational force and a force of friction proportional to velocity, then this mathematical model has the form

$$y'' + 2ay' + k^2 \sin y = 0, \quad y(0) = y_0, \quad y'(0) = 0, \tag{1}$$

where the unknown function y represents the angle made by the pendulum with the vertical as a function of the time t, primes denote differentiation with respect to t, a and k are given positive constants, and y_0 is a given constant (which may be assumed non-negative without loss of generality). Since the initial value problem (1) can not be solved explicitly, it is customary to restrict oneself to small oscillations; that is, to assume that y remains small and that sin y may be replaced by y. Thus, one considers instead of (1) the initial value problem

$$y'' + 2ay' + k^2 y = 0, \quad y(0) = y_0, \quad y'(0) = 0, \tag{2}$$

which is easily solved explicitly. Its solution is

$$y(t) = e^{-at}\left[y_0 \cos\sqrt{k^2 - a^2}\,t + \frac{ay_0}{\sqrt{k^2 - a^2}} \sin\sqrt{k^2 - a^2}\,t\right]$$

if the system is lightly damped, that is, if $a < k$. We shall assume $a < k$ throughout this paper, but the reader should encounter no serious difficulty in treating the case of heavy damping, $a \geqq k$, in a similar manner. If we let

$$k^2 - a^2 = \omega^2 > 0,\ A = ky_0/\omega,\ \delta = \arctan a/\omega,$$

then we can rewrite this solution as

$$y(t) = Ae^{-at}\cos(\omega t - \delta). \tag{3}$$

From this expression, we can derive physical information about the motion of the pendulum, for example that it oscillates with period $2\pi/\omega$ and exponentially decreasing amplitude.

However, there is no reason beyond intuition to believe that the solution (3) of the simplified problem (2) is a good approximation to the solution of the original problem (1). While the student might expect that further experience in solving differential equations would enable him to solve (1), it turns out that it is impossible to obtain an explicit solution. This news is not as unpleasant as one might think, since (1) is not the right problem. It should be remembered that the derivation of (1) neglects all forces other than gravity and friction, and involves various other simplifying assumptions. By rights, we should be trying to solve a problem of the form

$$y'' + 2ay' + k^2\sin y = f(t, y, y'),\quad y(0) = y_0,\quad y'(0) = 0, \tag{4}$$

where the term $f(t, y, y')$ includes all effects neglected in the derivation of (1). It is reasonable to assume that $f(t, y, y')$ is small in some sense, for otherwise (1) would not be a useful mathematical model for the motion of the pendulum. However, it can not be assumed that a precise expression for $f(t, y, y')$ is known. Thus it does not even make sense to ask for an explicit expression for the solution of (4). The proper question to ask is whether, for a given class of functions f which are small in some sense, the solution of (4) can be approximated by the solution of a problem such as (2) which is simple enough that it can be solved explicitly. The purpose of this paper is to provide, by quite elementary methods, an affirmative answer to this question. The methods, while directed at this very specific problem, indicate a few of the principal ideals in the qualitative theory of nonlinear differential equations.

2. Let $x(t)$ be the solution of the simplified initial value problem

$$x'' + 2ax' + k^2x = 0,\quad x(0) = y_0,\quad x'(0) = 0, \tag{5}$$

namely $x(t) = Ae^{-at}\cos(\omega t - \delta)$, where

$$k^2 - a^2 = \omega^2 > 0,\quad A = ky_0/\omega,\quad \delta = \arctan a/\omega.$$

Let $y(t)$ be the solution of

$$y'' + 2ay' + k^2y = p(t, y, y'),\quad y(0) = y_0,\quad y'(0) = 0. \tag{6}$$

We will assume that there exists a constant $M > 0$ such that

$$|p(t, y, y')| \leqq M(|y|^2 + |y'|^2) \tag{7}$$

for $t \geqq 0$ and for sufficiently small $|y|$ and $|y'|$. We remark that the problem (4), which would appear to be the proper one to consider rather than (6), is actually included in the form (6). Since $|\sin y - y| \leqq |y|^3$ for small $|y|$, we can write (4) in the form (6) with $p(t, y, y') = f(t, y, y') - k^2(\sin y - y)$. The function $f(t, y, y')$ obeys the same condition (7) as does $p(t, y, y')$.

Our goal is to show that $y(t)$ behaves in the same way as does $x(t)$ for large t. Since, as a crude approximation, we wish to show that $|y(t)|$ is no greater than a constant multiple of e^{-at}, we begin by makin the changes of variable $x = e^{-at}u$ in (5) and $y = e^{-at}v$ in (6). Then (5) becomes

$$u'' + \omega^2 u = 0, \quad u(0) = y_0, \quad u'(0) = ay_0 \tag{8}$$

with solution $u(t) = A\cos(\omega t - \delta)$. Also, (6) becomes

$$\begin{aligned} v'' + \omega^2 v &= e^{at}p(t, e^{-at}v, \quad (e^{-at}v)') \\ &= q(t, v, v'), \quad v(0) = y_0, \quad v'(0) = ay_0. \end{aligned} \tag{9}$$

Using (7) and $(e^{-at}v)' = e^{-at}(v' - av)$, we see that

$$\begin{aligned} |q(t, v, v')| &= |e^{at}p(t, e^{-at}v, (e^{-at}v)')| \\ &\leqq e^{at}M(|e^{-at}v|^2 + |e^{-at}(v' - av)|^2) \\ &\leqq e^{-at}M(|v|^2 + |v' - av|^2) \\ &\leqq Le^{-at}(|v|^2 + |v'|^2) \end{aligned} \tag{10}$$

for some constant $L > 0$.

THEOREM 1. *If* (10) *is satisfied, then there exists a constant* $B > 0$ *such that every solution* $v(t)$ *of the initial value problem* (9) *with* y_0 *sufficiently small satisfies*

$$|v(t)| \leqq B, \quad |v'(t)| \leqq B \tag{11}$$

for all $t \geqq 0$.

Proof: We consider (8) as a linear homogeneous differential equation and (9) as a non-homogeneous problem, to which we apply the variation of constants formula [**1**, Section 3.8]. Since the unknown function v appears in the non-homogeneous term in (9), we obtain an integral equation for $v(t)$ rather than an explicit formula; this integral equation is

$$v(t) = u(t) + \frac{1}{\omega}\int_0^t \sin\omega(t - s)\, q(s, v(s), v'(s))\, ds. \tag{12}$$

The equation (12) may be differentiated to yield

$$v'(t) = u'(t) + \int_0^t \cos\omega(t - s)\, q(s, v(s), v'(s))\, ds. \tag{13}$$

When we solve (8), we see from the explicit solution that $|u(t)| \leqq c$, $|u'(t)| \leqq c$ for $t \geqq 0$, where c is a constant which can be made arbitrarily small by making y_0 sufficiently small. In fact, $|u(t)| \leqq A = ky_0/\omega$, $|u'(t)| \leqq \omega A = ky_0$, and we can take $c = ky_0$ if $\omega \geqq 1$, $c = ky_0/\omega$ if $\omega < 1$. Now we estimate in (12) and (13) using (10); we obtain

$$\begin{aligned} |v(t)| &\leqq c + \frac{1}{\omega}\int_0^t |q(s,v(s),v'(s))|\,ds \\ &\leqq c + \frac{1}{\omega}\int_0^t Le^{-as}(|v(s)|^2 + |v'(s)|^2)\,ds, \\ |v'(t)| &\leqq c + \int_0^t |q(s,v(s),v'(s))|\,ds \\ &\leqq c + \int_0^t Le^{-as}(|v(s)|^2 + |v'(s)|^2)\,ds. \end{aligned} \tag{14}$$

We let $K = L$ if $\omega \geqq 1$, $K = L/\omega$ if $\omega < 1$ and add the two inequalities in (14), obtaining

$$\begin{aligned} |v(t)| + |v'(t)| &\leqq 2c + 2K\int_0^t e^{-as}(|v(s)|^2 + |v'(s)|^2)\,ds \\ &\leqq 2c + 2K\int_0^t e^{-as}(|v(s)| + |v'(s)|)^2\,ds. \end{aligned} \tag{15}$$

We are trying to show that $|v(t)|$ and $|v'(t)|$ are both bounded for $t \geqq 0$; to this end we let $r(t) = |v(t)| + |v'(t)|$, so that $r(t) \geqq 0$ and (15) becomes

$$r(t) \leqq 2c + 2K\int_0^t e^{-as}[r(s)]^2\,ds. \tag{16}$$

We point out that if for some $t_0 \geqq 0$ we have $r(t_0) = 0$, then $v(t_0) = v'(t_0) = 0$. Since (10) implies $q(t,0,0) = 0$ for all $t \geqq 0$, the identically zero function is a solution of the differential equation $v'' + \omega^2 r = q(t,v,v')$. By the uniqueness theorem for second order differential equations [1, Section 1.8], this implies $r(t) \equiv 0$, or $v(t) \equiv 0$. In this case, we must have $y_0 = 0$, and the solution $y(t)$ of (6) is identically zero. Thus if $r(t_0) = 0$ for some t_0, we are dealing with a trivial case, and we may assume $r(t) > 0$ for all $t \geqq 0$.

There is a standard type of argument in the theory of differential equations to show that a strictly positive function $r(t)$ which satisfies the integral inequality (16) is bounded. We let $R(t) = 2c + 2K\int_0^t e^{-as}[r(s)]^2 ds$, so that by (16), $0 < r(t) \leqq R(t)$. Also, $R(0) = 2c$, and

$$R'(t) = 2Ke^{-at}[r(t)]^2 \leqq 2Ke^{-at}[R(t)]^2. \tag{17}$$

Since $R(t) > 0$, we may divide (17) by $[R(t)]^2$, obtaining

$$\frac{R'(t)}{[R(t)]^2} \leqq 2Ke^{-at}. \tag{18}$$

We now integrate (18) from 0 to t, and obtain

$$\int_0^t \frac{R'(s)}{[R(s)]^2} ds \leqq 2K \int_0^t e^{-as} ds,$$

or

$$-\frac{1}{R(s)}\Bigg]_{s=0}^{s=t} \leqq \frac{2K}{a}(1-e^{-at}), \quad \frac{1}{2c} - \frac{1}{R(t)} \leqq \frac{2K}{a}(1-e^{-at}) \leqq \frac{2K}{a}.$$

From this, we see that

$$\frac{1}{R(t)} \geqq \frac{1}{2c} - \frac{2K}{a} = \frac{a-4cK}{2ac},$$

or $R(t) \leqq 2ac/(a-4cK)$, provided $c < a/4K$ (which can be achieved by taking y_0 small enough). We let $B = 2ac/(a-4cK)$, and then $R(t) \leqq B$ for $t \geqq 0$. Since $r(t) \leqq R(t)$, we have $r(t) = |v(t)| + |v'(t)| \leqq B$ for all $t \geqq 0$, which implies (11) and completes the proof of Theorem 1.

3. By returning to the variation of constants formulae (12) and (13), we may now obtain more precise information about the behavior of $v(t)$ and $v'(t)$ for large t.

THEOREM 2. *If* (10) *is satisfied, then for every solution* $v(t)$ *of* (9) *with* y_0 *sufficiently small there exist constants* $\hat{A}, \hat{\delta}, C > 0$ *such that* $v(t) = \hat{A}\cos(\omega t - \hat{\delta}) + h(t)$, *where* $|h(t)| \leqq Ce^{-at}$, $|h'(t)| \leqq Ce^{-at}$ *for* $t \geqq 0$.

Proof: We rewrite (12) and (13) as

$$\begin{aligned} v(t) = u(t) &+ \frac{1}{\omega}\int_0^\infty \sin\omega(t-s)\, q(s, v(s), v'(s))\, ds \\ &- \frac{1}{\omega}\int_0^\infty \sin\omega(t-s)\, q(s, v(s), v'(s))\, ds, \end{aligned} \tag{19}$$

$$\begin{aligned} v'(t) = u'(t) &+ \int_0^\infty \cos\omega(t-s)\, q(s, v(s), v'(s))\, ds \\ &- \int_t^\infty \cos\omega(t-s)\, q(s, v(s), v'(s))\, ds \end{aligned} \tag{20}$$

respectively. These formulae are valid if the infinite integrals converge. Since

$$\begin{aligned} |\sin\omega(t-s)\, q(s, v(s), v'(s))| &\leqq |q(s, v(s), v'(s))| \\ &\leqq Le^{-as}(|v(s)|^2 + |v'(s)|^2) \leqq 2LB^2 e^{-as}, \end{aligned} \tag{21}$$

using (10) and (11), the infinite integral in (19) converges. A similar argument proves the convergence of the infinite integral in (20).

Next, we observe that if we define

$$(22)\qquad \begin{aligned}\hat{u}(t) &= u(t) + \frac{1}{\omega}\int_0^\infty \sin\omega(t-s)\,q(s,v(s),v'(s))\,ds \\ &= u(t) + \sin\omega t\left[\frac{1}{\omega}\int_0^\infty q(s,v(s),v'(s))\cos\omega s\,ds\right] \\ &\quad - \cos\omega t\left[\frac{1}{\omega}\int_0^\infty q(s,v(s),v'(s))\sin\omega s\,ds\right],\end{aligned}$$

then $\hat{u}(t)$ is a linear combination of solutions of the linear homogeneous differential equation

$$(23)\qquad u'' + \omega^2 u = 0,$$

and is therefore itself a solution of (23). Now (19) becomes

$$(24)\qquad v(t) = \hat{u}(t) - \frac{1}{\omega}\int_t^\infty \sin\omega(t-s)\,q(s,v(s),z'(s))\,ds.$$

From the definition of $\hat{u}(t)$, it is easy to verify that (20) becomes

$$(25)\qquad v'(t) = \hat{u}'(t) - \int_t^\infty \cos\omega(t-s)\,q(s,v(s),v'(s))\,ds.$$

Using (21) and the fact that $\int_t^\infty e^{-as}ds = e^{-at}/a$, we see that we can write (24) and (25) as $v(t) = \hat{u}(t) + h(t)$, where $|h(t)| \leqq Ce^{-at}$, $|h'(t)| \leqq Ce^{-at}$ for all $t \geqq 0$ with some constant $C > 0$. Since every solution $\hat{u}(t)$ of (23) has the form $\hat{u}(t) = \hat{A}\cos(\omega t - \hat{\delta})$, the proof of Theorem 2 is complete.

Returning to the original variables $x = e^{-at}u$ and $y = e^{-at}v$ and applying Theorem 2, we can now give the desired result for the original problem.

THEOREM 3. *Let $y(t)$ be the solution of the initial value problem* (6), *where $a < k$ and $y_0 > 0$. Let $\omega^2 = k^2 - a^2$. Suppose that $p(t,y,y')$ satisfies* (7). *Then there exist constants $\hat{A}, \hat{\delta}, C > 0$ such that*

$$(26)\qquad y(t) = \hat{A}e^{-at}[\cos(\omega t - \hat{\delta}) + h(t)],$$

where $|h(t)| \leqq Ce^{-at}$, $|h'(t)| \leqq Ce^{-at}$ for all $t \geqq 0$.

4. The reader will observe that the formula (26) says that every solution of the "correct" initial value problem (6) with y_0 sufficiently small behaves like some solution of the idealized linear differential equation $y'' + 2ay' + k^2y = 0$. It does not, however, say that the solution of (6) behaves like the solution of the initial value problem (5), which is composed of the idealized differential equation $y'' + 2ay'$

$+ k^2y = 0$ together with the same initial conditions as those in (6). We recall that the solution of (5) is $x(t) = e^{-at}u(t)$, where

$$u(t) = A\cos(\omega t - \delta) = y_0 \cos \omega t + \frac{ay_0}{\omega} \sin \omega t.$$

The amplitude A is given by

$$A^2 = y_0^2 + \left(\frac{ay_0}{\omega}\right)^2, \tag{27}$$

and the phase angle δ is given by

$$\delta = \arctan a/\omega. \tag{28}$$

From Theorem 3, we see that the solution $y(t)$ of (6) is approximated by $e^{-at}\hat{u}(t)$, where $\hat{u}(t) = \hat{A}\cos(\omega t - \hat{\delta})$ is defined in (22). Before we can use the idealized problem (5) instead of the true problem (6) to make physical predictions about the motion of the simple pendulum, we must show that the amplitudes A and $\hat{A}$, and the phase angles δ and $\hat{\delta}$ are close together.

THEOREM 4. *If $c > 1$ is satisfied, then it is possible to make $\hat{A}$ arbitrarily close to A and $\hat{\delta}$ arbitrarily close to δ in* (26) *by choosing y_0 sufficiently small.*

Proof: In (22), if

$$\begin{aligned} d_1 &= \frac{1}{\omega} \int_0^\infty q(s, v(s), v'(s)) \cos \omega s \, ds, \\ d_2 &= -\frac{1}{\omega} \int_0^\infty q(s, v(s), v'(s)) \sin \omega s \, ds, \end{aligned}$$

we have

$$\begin{aligned} \hat{u}(t) &= u(t) + d_1 \sin \omega t + d_2 \cos \omega t \\ &= \left(\frac{ay_0}{\omega} + d_1\right) \sin \omega t + (y_0 + d_2) \cos \omega t. \end{aligned}$$

The amplitude $\hat{A}$ and phase angle $\hat{\delta}$ are given by

$$\hat{A}^2 = \left(\frac{ay_0}{\omega} + d_1\right)^2 + (y_0 + d_2)^2 = A^2 + \frac{2ay_0d_1}{\omega} + d_1^2 + 2d_2y_0 + d_2^2 \tag{29}$$

$$\hat{\delta} = \arctan \frac{(ay_0/\omega) + d_1}{y_0 + d_2}. \tag{30}$$

From (10), which was a consequence of (7), and $|v(t)| \leqq B$, $|v'(t)| \leqq B$, where $B = 2ac/(a - 4ck)$ and $c = ky_0$ if $\omega \geqq 1$, $c = ky_0/\omega$ if $\omega < 1$, we have

$$\int_0^\infty |q(s,v(s),v'(s))|\,ds \leqq 2LB^2 \int_0^\infty e^{-as}ds = \frac{2LB^2}{a}.$$

From this it follows that $|d_1| \leqq 2LB^2/a$, $|d_2| \leqq 2LB^2/a$. Since B can be made arbitrarily small by making y_0 sufficiently small, we now see from (29) that $\hat{A}^2 - A^2$ can be made arbitrarily small and from (30) that $\tan\hat{\delta} - \tan\delta$, and hence $\hat{\delta} - \delta$, can be made arbitrarily small by making y_0 sufficiently small. This completes the proof of Theorem 4.

The reader should note that the true amplitude $\hat{A}$ and phase angle $\hat{\delta}$ given in (29) and (30) respectively can not be calculated exactly, because the numbers d_1 and d_2 depend on the unknown function q and the unknown solution v. We can only approximate them, and we can do even this only for sufficiently small y_0. For practical applications it is extremely important to know how small y_0 must be for our results to be applicable. Unfortunately, our approach yields no information about this problem. This gap is more or less characteristic of non-linear differential equations, and suggests a large class of largely unsolved problems.

This work was supported by the National Science Foundation, Contract No. GP-28267.

Reference

1. F. Brauer and J. A. Nohel, Ordinary Differential Equations: A First Course, Benjamin, New York, 1967.

TRUTH WITH RESPECT TO AN ULTRAFILTER OR HOW TO MAKE INTUITION RIGOROUS

D. H. VAN OSDOL, University of New Hampshire

Dedicated to Harold B. Hanes

Introduction. Our purpose in this article is to give a very concrete, simple exposition of some of the ideas of Abraham Robinson [**2**]. Our approach follows the outline sketched by Professor Takahashi at the Université de Montréal in the summer of 1970. This consists of constructing a particular non-standard model of the real numbers in which our intuition seems to work. The philosophy is then to take a conjecture about the reals, interpret and prove it in the non-standard model, and then conclude that the conjecture holds for the real numbers. We introduce ultrafilters in Section 1, construct the non-standard model in Section 2, and give non-standard proofs of some calculus theorems in Section 3.

Donovan Van Osdol received his Illinois Ph. D. in 1969 under John Gray and Michael Barr. Since then he has taught at Wilkes College and the University of New Hampshire. His research interests are homological algebra and category theory. *Editor.*

1. Ultrafilters. Let N, Z, Q, R be respectively the sets of natural numbers, integers, rational numbers, and real numbers as developed, say, in [**1**]. Consider $\mathscr{F}$, a collection of subsets of N, defined as follows: Let $\mathscr{F} = \{S \subseteq N | N \sim S$ is finite$\}$. This collection of subsets of N is a **filter** on N, that is, $\mathscr{F}$ satisfies the following three properties:

(i) $\varnothing \notin \mathscr{F}$.

(ii) If $S_1, S_2 \in \mathscr{F}$, then $S_1 \cap S_2 \in \mathscr{F}$.

(iii) If $S \in \mathscr{F}$ and $S \subseteq T \subseteq N$, then $T \in \mathscr{F}$.

If we say that a function f with domain N has property P whenever $\{n \in N | f(n)$ has property $P\} \in \mathscr{F}$, then the filter properties of $\mathscr{F}$ translate into the following logical properties:

(i) If $f(n)$ has property P for no $n \in N$ then f does not have property P.

(ii) If f has property P and property Q then f has property $P \wedge Q$ (= the logical conjunction of P and Q).

(iii) If f has property P and P implies Q, then f has property Q.

Since we are soon going to make precisely such a definition, we are happy to have the above logical properties.

There is a disappointment in this approach, however. Take $f(n) = 1$ if n is odd, $f(n) = -1$ if n is even. We would like to be able to say whether $f \geqq 0$ or $f < 0$; but clearly it is neither. This violates the basic law of logic which says that given a property P and an entity f, either f has property P or f has property "not P" (and not both). Clearly the way to avoid this problem is to have a filter $\mathscr{U}$ in which, for each $S \subseteq N$, either $S \in \mathscr{U}$ or $N \sim S \in \mathscr{U}$. Such a filter is called an **ultrafilter**. We are now going to prove that there is at least one ultrafilter $\mathscr{U}$ on N which contains $\mathscr{F}$ as a subset.

Intuitively, an ultrafilter containing $\mathscr{F}$ would have to be a collection of subsets of N which contains as many sets as possible (consistent with it being a filter containing $\mathscr{F}$). Consequently, we let $F = \{$filters $\mathscr{F}'$ on $N | \mathscr{F}' \supseteq \mathscr{F}\}$ and try to apply Zorn's lemma. We take the order relation on F to be containment, and notice that $F \neq \varnothing$ (since $\mathscr{F} \in F$). Also, any chain in F has an upper bound in F, namely the union over the chain. Hence Zorn's lemma implies the existence of at least one maximal member, $\mathscr{U}$, of F. We claim that $\mathscr{U}$ is an ultrafilter on N. If not, let $\varnothing \neq S \subseteq N$ satisfy $S \notin \mathscr{U}$ and $N \sim S \notin \mathscr{U}$. There must exist a $T \in \mathscr{U}$ such that $S \cap T = \varnothing$ because otherwise $\mathscr{U}' = \{X \subseteq N | X \supseteq S \cap T$ for some $T \in \mathscr{U}\} \in F$ and $\mathscr{U}' \supset \mathscr{U}$, contradicting the maximality of $\mathscr{U}$. Similarly there is a $T' \in \mathscr{U}$ such that $(N \sim S) \cap T' = \varnothing$. But this is absurd, because we then have

$$\begin{aligned} \varnothing &= ((N \sim S) \cap T') \supseteq ((N \sim S) \cap T \cap T') \cup (S \cap T \cap T') \\ &= T \cap T' \in \mathscr{U}. \end{aligned}$$

Thus $\mathscr{U}$ is an ultrafilter.

For the rest of this paper, the symbol $\mathscr{U}$ will represent an ultrafilter on N such that $\mathscr{U} \supset \mathscr{F}$.

2. The Ultrapower $^*\boldsymbol{R}$. It is now possible to define a specific "non-standard" model of the real numbers. Let $\boldsymbol{R}^{\boldsymbol{N}}$ be the set of all functions from $\boldsymbol{N}$ to $\boldsymbol{R}$, and interpret truth in this set with respect to $\mathscr{U}$. For example, given $f, g \in \boldsymbol{R}^{\boldsymbol{N}}$ say that f "$=$" g if $\{n \in \boldsymbol{N} | f(n) = g(n)\} \in \mathscr{U}$. It is easy to show that "$=$" is a relation on $\boldsymbol{R}^{\boldsymbol{N}}$ which is reflexive because $\boldsymbol{N} \in \mathscr{U}$, symmetric because $=$ is symmetric on $\boldsymbol{R}$, and transitive because of condition (ii) for a filter. Let $^*\boldsymbol{R}$ be the set of equivalence classes of $\boldsymbol{R}^{\boldsymbol{N}}$ with respect to "$=$", and write $\langle f \rangle$ for the element of $^*\boldsymbol{R}$ which represents $f \in \boldsymbol{R}^{\boldsymbol{N}}$. Thus $\langle f \rangle = \langle g \rangle$ in $^*\boldsymbol{R}$ if and only if $\{n \in \boldsymbol{N} | f(n) = g(n)\} \in \mathscr{U}$. This construction of $^*\boldsymbol{R}$ is called an *ultrapower*, presumably because it is arrived at by taking a cartesian power and then reducing modulo an ultrafilter. As often happens in mathematics, the crucial importance of the ultrafilter in defining an ultrapower is completely ignored in the notation for it.

Many properties of $\boldsymbol{R}$ are also inherited by $^*\boldsymbol{R}$. For example:

(i) $^*\boldsymbol{R}$ is a field. Define $\langle f \rangle + \langle g \rangle = \langle f + g \rangle$ and $\langle f \rangle \langle g \rangle = \langle fg \rangle$: that these operations are well-defined depends on property (ii) of a filter. For any $x \in \boldsymbol{R}$ let x: $\boldsymbol{N} \to \boldsymbol{R}$ be given by $x(n) = x$ for all $n \in \boldsymbol{N}$. Then the additive identity of $^*\boldsymbol{R}$ is $\langle 0 \rangle$, the multiplicative identity is $\langle 1 \rangle$, and the additive inverse of $\langle f \rangle$ is $\langle -f \rangle$. Multiplicative inverses are a bit more tricky: if $\langle g \rangle \neq \langle 0 \rangle$ then $\{n \in \boldsymbol{N} | g(n) = 0\} \notin \mathscr{U}$, and since $\mathscr{U}$ is an ultrafilter, $\{n \in \boldsymbol{N} | g(n) \neq 0\} \in \mathscr{U}$. If we define $h: \boldsymbol{N} \to \boldsymbol{R}$ by

$$h(n) = \begin{cases} g(n)^{-1} & \text{if} \quad g(n) \neq 0 \\ 0 & \text{if} \quad g(n) = 0, \end{cases}$$

then $\langle g \rangle^{-1} = \langle h \rangle$ because $\{n \in \boldsymbol{N} | g(n)h(n) = 1\} = \{n \in \boldsymbol{N} | g(n) \neq 0\} \in \mathscr{U}$. This implies $\langle g \rangle \langle h \rangle = \langle 1 \rangle$. The field axioms for $^*\boldsymbol{R}$ are now easily verified (because they hold for $\boldsymbol{R}$).

(ii) $^*\boldsymbol{R}$ is an ordered field. Define $\langle f \rangle \leqq \langle g \rangle$ if $\{n \in \boldsymbol{N} | f(n) \leqq g(n)\} \in \mathscr{U}$. This is a well-defined relation because if $\langle f \rangle = \langle f' \rangle$ and $\langle g \rangle = \langle g' \rangle$, then let $F = \{n \in \boldsymbol{N} | f(n) = f'(n)\} \in \mathscr{U}$, $G = \{n \in \boldsymbol{N} \mid g(n) = g'(n)\} \in \mathscr{U}$, and $L = \{n \in \boldsymbol{N} \mid f(n) \leqq g(n)\} \in \mathscr{U}$. Since $F \cap G \cap L \subseteq \{n \in \boldsymbol{N} | f'(n) \leqq g'(n)\}$, properties (ii) and (iii) of a filter combine to prove that $\langle f' \rangle \leqq \langle g' \rangle$. Of course $\langle f \rangle \leqq \langle g \rangle$ and $\langle f \rangle \neq \langle g \rangle$ if and only if $\{n \in \boldsymbol{N} | f(n) < g(n)\} \in \mathscr{U}$, for which situation we write $\langle f \rangle < \langle g \rangle$. The compatibility of $<$ with addition and multiplication is easy to verify, so we turn our attention to trichotomy. That is, given $\langle f \rangle, \langle g \rangle \in {}^*\boldsymbol{R}$ precisely one of: $\langle f \rangle = \langle g \rangle$; $\langle f \rangle < \langle g \rangle$; $\langle g \rangle < \langle f \rangle$ is true. If we let $E = \{n \in \boldsymbol{N} | f(n) = g(n)\}$, $L = \{n \in \boldsymbol{N} | f(n) < g(n)\}$, and $G = \{n \in \boldsymbol{N} \mid g(n) < f(n)\}$, then $E \cup L \cup G = \boldsymbol{N}$ because trichotomy holds in $\boldsymbol{R}$. For the same reason $E \cap L = E \cap G = L \cap G = \varnothing$. Since $\mathscr{U}$ is an ultrafilter, at least one of E,L,G must be an element of $\mathscr{U}$ (if $E \in \mathscr{U}$, fine; if not, its complement $L \cup G \in \mathscr{U}$; now if $L \in \mathscr{U}$, good, but if not then its complement $E \cup G \in \mathscr{U}$; hence $G = (L \cup G) \cap (E \cup G) \in \mathscr{U}$. On the other hand, no two of E,L,G can be in $\mathscr{U}$, because then their empty intersection would be in $\mathscr{U}$, contradicting property (i) of a filter. Hence, precisely one of E,L,G is in $\mathscr{U}$, and this is equivalent to trichotomy.

(iii) $\boldsymbol{R}$ is embedded in $^*\boldsymbol{R}$ as an ordered proper subfield. Define $i: \boldsymbol{R} \to {}^*\boldsymbol{R}$ by $i(x) = \langle x \rangle$; then i is a field homomorphism, and hence an embedding, which preserves order. However, i is not onto, for let $f: \boldsymbol{N} \to \boldsymbol{R}$ be defined by $f(n) = n$ for all $n \in \boldsymbol{N}$. There is no $x \in \boldsymbol{R}$ such that $i(x) = \langle f \rangle$ because if there were, then $\{n \in \boldsymbol{N} \mid f(n) = x\} \in \mathscr{U}$ would be a non-empty set. It would follow that $x \in \boldsymbol{N}$ and $\{x\} = \{n \in \boldsymbol{N} \mid f(n) = x\} \in \mathscr{U}$, which is impossible because $\mathscr{U}$ is an ultrafilter and $\boldsymbol{N} \sim \{x\} \in \mathscr{F} \subseteq \mathscr{U}$.

We conclude that $^*\boldsymbol{R}$ does not inherit completeness from $\boldsymbol{R}$, because $\boldsymbol{R}$ is the "only" complete ordered field. Our sacrifice of completeness is compensated for, however, by the existence of "infinitely large" and "infinitely small" elements in $^*\boldsymbol{R}$. Given $\langle f \rangle \in {}^*\boldsymbol{R}$ define $|\langle f \rangle| = \langle h \rangle$, where $h(n) = |f(n)|$. We say that $\langle f \rangle \in {}^*\boldsymbol{R}$ is **infinitely** large if $i(x) \leqq |\langle f \rangle|$ for each $x \in \boldsymbol{R}$, and is **infinitesimal** or **infinitely small** if $i(0) < |\langle f \rangle| \leqq i(x)$ for each $x \in \boldsymbol{R}$ such that $0 < x$. The existence of infinitely large and small elements in $^*\boldsymbol{R}$ is best demonstrated by producing one of each. If $f(n) = n$, then $\langle f \rangle$ is infinitely large because for each $x \in \boldsymbol{R}$, $\{n \in \boldsymbol{N} \mid f(n) > x\} \supseteq \{m, m+1, m+2, \cdots\} \in \mathscr{F} \subseteq \mathscr{U}$, where m is some natural number with $x \leqq m$. Similarly, if $g(n) = 1/n$, then $\langle g \rangle$ is infinitesimal (and $\langle g \rangle \neq i(0)$).

Another interesting aspect of $^*\boldsymbol{R}$ is that if we take the bounded elements (those which are not infinitely large) and identify those which are infinitely close together, then we end up with a field which is isomorphic to $\boldsymbol{R}$. More precisely, let $B = \{\langle f \rangle \in {}^*\boldsymbol{R} \mid \text{there is } x \in \boldsymbol{R} \text{ such that } |\langle f \rangle| \leqq i(x)\}$ and let $I = \{\langle f \rangle \in {}^*\boldsymbol{R} \mid \langle f \rangle \text{ is infinitesimal}\}$. Then I is the unique maximal ideal of the ring B, and $\boldsymbol{R} \cong B/I$. Clearly B is a ring and I is a subring. Given $\langle f \rangle \in B$ and $\langle g \rangle \in I$, let $|\langle f \rangle| \leqq i(x)$ with $x > 0$ and let $y > 0$ be arbitrary; then

$$\{n \in \boldsymbol{N} \mid |f(n)g(n)| \leqq y\} \supseteq$$
$$\{n \in \boldsymbol{N} \mid |f(n)| \leqq x\} \cap \{n \in \boldsymbol{N} \mid |g(n)| \leqq y/x\} \in \mathscr{U},$$

and hence $\langle f \rangle \langle g \rangle \in I$.

This shows I is an ideal of B. Any ideal J of B which contains an $\langle f \rangle \in B \sim I$ must be equal to B. This is because for each $\langle f \rangle \in B \sim I$, there exists $x \in \boldsymbol{R}$ such that $\{n \in \boldsymbol{N} \mid |f(n)| \geqq x\} \in \mathscr{U}$. Hence $\{n \in \boldsymbol{N} \mid |f(n)^{-1}| \leqq 1/x\} \in \mathscr{U}$, or equivalently $\langle f \rangle^{-1} \in B$. Thus $i(1) = \langle f \rangle \langle f \rangle^{-1} \in J$. It follows that I is the unique maximal ideal in B, and that B/I is a field. The mapping $\phi: \boldsymbol{R} \to B/I$ defined by $\phi(x) = i(x) + I$ is a homomorphism and will be an isomorphism if and only if it is onto. To show this, let $\langle f \rangle \in B$, $L = \{x \in \boldsymbol{R} \mid i(x) < \langle f \rangle\}$, and $U = \{x \in \boldsymbol{R} \mid \langle f \rangle \leqq i(x)\}$. The pair L, U forms a Dedekind cut for $\boldsymbol{R}$ (since $\mathscr{U}$ is an ultrafilter) and hence there is a unique $y \in \boldsymbol{R}$ such that if $x < y$ then $x \in L$ and if $y < x$ then $x \in U$. We shall show that $\langle f \rangle + I = \phi(y)$ by demonstrating that $\langle f \rangle - i(y) \in I$. Let $z > 0$ be given: then $y + z \in U$ and $y - z \in L$. Then $\{n \in \boldsymbol{N} \mid f(n) < y + z\} \in \mathscr{U}$ and $\{n \in \boldsymbol{N} \mid y - z < f(n)\} \in \mathscr{U}$, which implies $\{n \in \boldsymbol{N} \mid |f(n) - y| < z\} \in \mathscr{U}$. Thus $\langle f \rangle - i(y) \in I$, ϕ is onto, and $\boldsymbol{R} \cong B/I$.

We end this section with two more definitions. Given $X \subseteq \boldsymbol{R}$, let $^*X = \{\langle f \rangle \in {}^*\boldsymbol{R} \mid \{n \in \boldsymbol{N} \mid f(n) \in X\} \in \mathscr{U}\}$. Given $\alpha: X \to \boldsymbol{R}$, define $^*\alpha: {}^*X \to {}^*\boldsymbol{R}$ by the rule $^*\alpha(\langle f \rangle) = \langle g \rangle$ where

$$g(n) = \begin{cases} \alpha(f(n)) & \text{if } f(n) \in X \\ 0 & \text{if } f(n) \in X. \end{cases}$$

These are somewhat strange definitions (in the sense that $*X$ "should be" $\{\langle f\rangle \in *R \mid f(n) \in X \text{ for all } n \in N\}$), but they are consistent with our general program of interpreting truth relative to $\mathscr{U}$. They also turn out to be most useful.

3. Calculus Theorems. It is our purpose in this section to show the interplay between intuition and $*R$ by proving some calculus theorems. For example, when one thinks of a sequence converging to a number, he imagines that if he goes out infinitely far in the sequence then he will be infinitely near to the limit. Freshman calculus students try to make this intuition work for themselves when they substitute infinity into the general term of a sequence and then try to read off the limit from the expression which they get. We now see that in some sense this is the correct thing to do.

THEOREM 1. *A sequence $S: N \to R$ converges to $r \in R$ if and only if for each infinitely large $\langle v\rangle \in *N$ we have $*S(\langle v\rangle) - i(r)$ infinitesimal.*

Proof. If S converges to r, let $x > 0$ be given and $\langle v\rangle \in *N$ infinitely large. There is $m \in N$ such that $|S(n) - r| < x$ for each $n > m$, so let

$$F = \{n \in N \mid v(n) > m\} \cap \{n \in N \mid v(n) \in N\} \in \mathscr{U}.$$

Thus $\{n \in N \mid |S(v(n)) - r| < x\} \supseteq F$, hence is itself in $\mathscr{U}$, and $*S(\langle v\rangle) - i(r)$ is infinitesimal. Conversely, suppose S does not converge to r. Then there exists $x > 0$ such that for each $m \in N$ there is an $n \geqq m$ with $|S(n) - r| \geqq x$. For each $m \in N$ define $v(m) = n$ where $n \geqq m$ and $|S(n) - r| \geqq x$. It follows that $\langle v\rangle \in *N$ is infinitely large and $\{n \in N \mid |S(v(n)) - r| \geqq x\} = N \in \mathscr{U}$. Thus $|*S(\langle v\rangle) - i(r)|$ is not infinitely small.

Example. We shall show that $\sum_{i=1}^{\infty} 1/i(i+1) = 1$. Let $S(n) = \sum_{i=1}^{n} 1/i(i+1)$. For any infinitely large $\langle v\rangle \in *N$ we have

$$\begin{aligned} *S(\langle v\rangle) - 1 &= \sum_{i=1}^{\langle v\rangle} \frac{1}{i(i+1)} - 1 \\ &= \sum_{i=1}^{\langle v\rangle} \left(\frac{1}{i} - \frac{1}{i+1}\right) - 1 \\ &= \sum_{i=1}^{\langle v\rangle} \frac{1}{i} - \sum_{i=1}^{\langle v\rangle} \frac{1}{i+1} - 1 \\ &= -\frac{1}{\langle v\rangle + 1}, \end{aligned}$$

which is infinitesimal.

We now turn to continuity. We feel that a function α is continuous at x if whenever we take a point infinitely close to x then α of it is infinitely close to $\alpha(x)$. We simply need to interpret this intuition in $*R$ in order to get a theorem.

THEOREM 2. *A function* $\alpha: X \to \mathbf{R}$ *is continuous at* $x \in X$ *if and only if whenever* $\langle f \rangle \in {}^*X$ *is such that* $i(x) - \langle f \rangle$ *is infinitesimal also* ${}^*\alpha(i(x)) - {}^*\alpha(\langle f \rangle)$ *is infinitesimal.*

Proof. If α is continuous at $x \in X$ then for each $a > 0$ there exists $b > 0$ such that $|\alpha(x) - \alpha(y)| < a$ for each $y \in X$ with $|x - y| < b$. Now $i(x) - \langle f \rangle$ infinitesimal implies

$$\{n \in \mathbf{N} |\ |x - f(n)| < b\} \cap \{n \in \mathbf{N} | f(n) \in X\} \in \mathscr{U}.$$

But $\{n \in \mathbf{N} | f(n) \in X \text{ and } |\alpha(x) - \alpha(f(n))| < a\}$ contains this set, hence is itself in $\mathscr{U}$, and ${}^*\alpha(i(x)) - {}^*\alpha(\langle f \rangle)$ is infinitely small. Conversely, if α is not continuous at $x \in X$ then there exists $a > 0$ such that for each $b > 0$ there is a $y \in X$ with $|x - y| < b$ and $|\alpha(x) - \alpha(y)| \geqq a$. Define $f: \mathbf{N} \to X$ by $f(n) = y_n$, where $|x - y_n| < 1/n$ and $|\alpha(x) - \alpha(y_n)| \geqq a$. Then $\langle f \rangle \in {}^*X$ and $\{n \in \mathbf{N} |\ |\alpha(x) - \alpha(f(n))| \geqq a\} = \mathbf{N} \in \mathscr{U}$, so that ${}^*\alpha(i(x)) - {}^*\alpha(\langle f \rangle)$ is not infinitesimal. However, for an arbitrary $c > 0$, $\{n \in \mathbf{N} |\ |x - f(n)| < c\} \in \mathscr{U}$ because it contains $\{m, m+1, m+2, \cdots\} \in \mathscr{F} \subseteq \mathscr{U}$, where $m \in \mathbf{N}$ and $m > 1/c$. Thus $i(x) - \langle f \rangle$ is infinitely small.

Example. The function $\alpha: (0, \infty) \to \mathbf{R}$ defined by $\alpha(x) = 1/x$ is continuous. For given any $x > 0$ and any $\langle f \rangle \in {}^*(0, \infty)$ such that $\{n \in \mathbf{N} | f(n) > 0$ and $|x - f(n)| < c\} \in \mathscr{U}$ for any $c > 0$, we have in particular $\{n \in \mathbf{N} | f(n) > 0$ and $|x - f(n)| < \frac{1}{2}x\} \in \mathscr{U}$ and $\{n \in \mathbf{N} | f(n) > 0$ and $|x - f(n)| < \frac{1}{2}cx^2\} \in \mathscr{U}$. Thus $\{n \in \mathbf{N} | f(n) > 0$ and $|\alpha(x) - \alpha(f(n))| < c\} \in \mathscr{U}$, that is, ${}^*\alpha(i(x)) - {}^*\alpha(\langle f \rangle)$ is infinitesimal.

A mapping should be uniformly continuous if it takes infinitely near points to infinitely near points. Again, this is true in ${}^*\mathbf{R}$.

THEOREM 3. *A function* $\alpha: X \to \mathbf{R}$ *is uniformly continuous if and only if for each* $\langle f \rangle, \langle g \rangle \in {}^*X$ *such that* $\langle f \rangle - \langle g \rangle$ *is infinitesimal also* ${}^*\alpha(\langle f \rangle) - {}^*\alpha(\langle g \rangle)$ *is infinitesimal.*

Proof. If α is uniformly continuous, then for each $a > 0$ there is $b > 0$ such that $x, y \in X$ and $|x - y| < b$ implies $|\alpha(x) - \alpha(y)| < a$. Thus if $\langle f \rangle, \langle g \rangle \in {}^*X$ and $\langle f \rangle - \langle g \rangle$ is infinitesimal, then $\{n \in \mathbf{N} | f(n), g(n) \in X$ and $|\alpha(f(n)) - \alpha(g(n))| < a\} \supseteq \{n \in \mathbf{N} | f(n), g(n) \in X$ and $|f(n) - g(n)| < b\} \in \mathscr{U}$, and ${}^*\alpha(\langle f \rangle) - {}^*\alpha(\langle g \rangle)$ is infinitesimal. Conversely, suppose α is not uniformly continuous. Then there exists $a > 0$ such that for each $n \in \mathbf{N}$, we can find $x_n, y_n \in X$ with $|x_n - y_n| < 1/n$ and $|\alpha(x_n) - \alpha(y_n)| \geqq a$. Define $f(n) = x_n$, $g(n) = y_n$. By arguments which are now familiar, $\langle f \rangle - \langle g \rangle$ is infinitely small but ${}^*\alpha(\langle f \rangle) - {}^*\alpha(\langle g \rangle)$ is not.

Example. The function $\alpha: (0, \infty) \to \mathbf{R}$ defined by $\alpha(x) = 1/x$ is not uniformly continuous. For let $f(n) = 1/n$, $g(n) = 1/(n+1)$; then $\langle f \rangle - \langle g \rangle$ is infinitesimal but ${}^*\alpha(\langle f \rangle) - {}^*\alpha(\langle g \rangle) = i(-1)$.

Shifting attention now to topology, a set "should be" closed if any point which is infinitely near some point in the set is itself in the set.

THEOREM 4. *A set $X \subseteq \boldsymbol{R}$ is closed if and only if for each $\langle f \rangle \in {}^*X$ and each $y \in \boldsymbol{R}$ such that $\langle f \rangle - i(y)$ is infinitely small, it follows that $y \in X$.*

Proof. Suppose X is closed, $\langle f \rangle \in {}^*X$, $y \in \boldsymbol{R}$, $\langle f \rangle - i(y)$ is infinitesimal, and let $U = (y - a, y + a)$ be a basic neighborhood of y. Then $\{n \in N \mid f(n) \in X$ and $|f(n) - y| < a\} \in \mathscr{U}$ (since $\langle f \rangle \in {}^*X$ and $\langle f \rangle - i(y)$ is infinitesimal), and in particular, is not empty. Hence there is an $n \in N$ such that $f(n) \in X \cap U$. Since X is closed and every neighborhood of y meets X, $y \in X$. Conversely, if X is not closed then there is a $y \in \boldsymbol{R}$ such that every neighborhood of y meets X, but $y \notin X$. For each $n \in N$ let $x_n \in (y - 1/n,\ y + 1/n) \cap X$ and define $f(n) = x_n$. Then $\langle f \rangle \in {}^*X$ and $\langle f \rangle - i(y)$ is infinitesimal, but $y \notin X$.

THEOREM 5. *A set $X \subseteq \boldsymbol{R}$ is open if and only if for each $x \in X$ and each $\langle f \rangle \in {}^*\boldsymbol{R}$ such that $\langle f \rangle - i(x)$ is infinitesimal then necessarily $\langle f \rangle \in {}^*X$.*

Proof. A set is closed if and only if its complement is open. Use Theorem 4.

The term "compact" connotes a collection which is closely packed or knit together. Our next theorem says essentially that.

THEOREM 6. *A set $X \subseteq \boldsymbol{R}$ is compact if and only if for each $\langle f \rangle \in {}^*X$ there exists a unique $x \in X$ such that $\langle f \rangle - i(x)$ is infinitesimal.*

Proof. Suppose X is compact and $\langle f \rangle \in {}^*X$, but $\langle f \rangle - i(x)$ is not infinitesimal for any $x \in X$. Then for each $x \in X$ there is an $a_x > 0$ such that $F_x = \{n \in N \mid f(n) \in X$ and $|f(n) - x| \geqq a_x\} \in \mathscr{U}$. The open intervals $(x - a_x, x + a_x)$ cover X, hence there is a finite subcover $(x_1 - a_{x_1}, x_1 + a_{x_1}), \cdots, (x_m - a_{x_m}, x_m + a_{x_m})$. Thus $\varnothing = F_{x_1} \cap \cdots \cap F_{x_m} \in \mathscr{U}$, a contradiction. It follows that there is at least one $x \in X$ with $\langle f \rangle - i(x)$ infinitesimal. If $i(y) - \langle f \rangle$ is also infinitely small for some $y \in X$ then $i(y) - i(x) = i(y) - \langle f \rangle + \langle f \rangle - i(x)$ is infinitesimal, from which we infer that $x = y$. Conversely, suppose X is not compact and $\mathscr{B}$ is an open cover of X possessing no finite subcover. For each $x \in X$, x is in some $B \in \mathscr{B}$, and we can find an open interval with rational endpoints which contains x and is contained in B. If we produce such an open interval for each $x \in X$, we get a countable set of open intervals. Moreover, we know that to each such open interval there corresponds a $B \in \mathscr{B}$ containing it. If we let $\mathscr{C}$ be the set of B's in $\mathscr{B}$ arising in this way then $\mathscr{C} \subseteq \mathscr{B}$ and $\mathscr{C}$ is a countable open cover of X. Unfortunately, $\mathscr{C}$ may still have "too many" sets in it to suit us. Let $\mathscr{C}_0 = \mathscr{C} = \{C_1, C_2, C_3, \cdots\}$. We first throw out all sets in $\mathscr{C}_0$ which are contained in C_1, and let $\mathscr{C}_1 \subseteq \mathscr{C}_0$ be the sets which remain. Let m_1 be the smallest $n > 1$ such that $C_n \in \mathscr{C}_1$ ($\mathscr{C}_1$ contains more than just C_1 because C_1 does not cover X). Delete from $\mathscr{C}_1$ all sets which are contained in $C_1 \cup C_{m_1}$, and call what is left $\mathscr{C}_2 \subseteq \mathscr{C}_1$. Let m_2 be the smallest $n > m_1$ such that $C_n \in \mathscr{C}_2$ (again such n exist because $\{C_1, C_{m_1}\}$ does not cover X) and delete from $\mathscr{C}_2$ all sets contained in $C_1 \cup C_{m_1} \cup C_{m_2}$. Let $\mathscr{C}_3 \subseteq \mathscr{C}_2$ be what is left, and continue inductively. Write $D_1 = C_1$ and $D_{n+1} = C_m$ for $n \in N$. Then $\mathscr{D} = \{D_n \mid n \in N\} \subseteq \mathscr{B}$

is a cover of X, and $D_n \sim (\bigcup_{i=1}^{n-1} D_i) \neq \varnothing$ for each $n \in N$. We pick $x_1 \in D_1$ and $x_n \in D_n \sim (\bigcup_{i=1}^{n-1} D_i)$ for each $n \geqq 2$. Define $f(n) = x_n$; then $\langle f \rangle \in {}^*X$ but $\langle f \rangle \notin {}^*D_m$ for any $m \in N$, because $\{n \in N \mid f(n) \notin D_m\} \supseteq \{m+1, m+2, \cdots\} \in \mathscr{U}$. For any $x \in X$, $x \in D_n$ for some $n \in N$ (because $\mathscr{D}$ covers X), so that by Theorem 5 we cannot have $\langle f \rangle - i(x)$ infinitesimal.

We now conclude by giving non-standard proofs of three theorems whose statements contain no non-standard terms.

THEOREM 7. *The continuous image of a compact set is compact.*

Proof. Let $X \subset \boldsymbol{R}$ be a compact set, $\alpha: X \to \boldsymbol{R}$ continuous, and $\langle f \rangle \in {}^*[\alpha(X)]$. For each $n \in \{n \in N \mid f(n) \in \alpha(X)\} \in \mathscr{U}$ pick an $x_n \in X$ such that $\alpha(x_n) = f(n)$, and define $g(n) = x_n$. For $n \notin \{n \in N \mid f(n) \in \alpha(X)\}$ define $g(n) = 0$. Then $\langle g \rangle \in {}^*X$, and there is a unique $x \in X$ such that $\langle g \rangle - i(x)$ is infinitesimal. Since α is continuous, ${}^*\alpha(\langle g \rangle) - {}^*\alpha(i(x))$ is infinitesimal (Theorem 2). Thus ${}^*\alpha(\langle g \rangle) - {}^*\alpha(i(x)) = \langle f \rangle - i(\alpha(x))$ is infinitesimal, and we are finished because of Theorem 6.

THEOREM 8. *A set $X \subseteq \boldsymbol{R}$ is compact if and only if it is closed and bounded.*

Proof. It is easy to verify that X is bounded if and only if ${}^*X$ contains no infinitely large elements. Thus, if X is compact, then it is bounded. For given $\langle f \rangle \in {}^*X$ let $x \in X$ be such that $\langle f \rangle - i(x)$ is infinitesimal. We have $\{n \in N \mid |f(n) - x| < 1\} \in \mathscr{U}$ and therefore $\{n \in N \mid x - 1 < f(n) < x + 1\} \in \mathscr{U}$, so that $\langle f \rangle$ is not infinitely large. Moreover, if X is compact, then it is closed. For given $\langle f \rangle \in {}^*X$ and $\langle f \rangle - i(y)$ infinitesimal for some $y \in \boldsymbol{R}$, let $x \in X$ be such that $i(x) - \langle f \rangle$ is infinitesimal (Theorem 6). Then $i(x) - i(y) = i(x) - \langle f \rangle + \langle f \rangle - i(y)$ is the sum of two infinitesimals, hence is itself infinitesimal, and $x = y$. Thus $y \in X$, and we have verified Theorem 2. Conversely, suppose X is closed and bounded, and let $\langle f \rangle \in {}^*X$. Then $\langle f \rangle$ is not infinitely large, so there exists $r \in \boldsymbol{R}$ such that $\{n \in N \mid |f(n)| \leqq r\} \in \mathscr{U}$. If $\langle f \rangle = i(x)$ for some $x \in \boldsymbol{R}$ we are done. If not, the sets $A = \{x \in \boldsymbol{R} \mid i(x) < \langle f \rangle\}$ and $B = \{x \in \boldsymbol{R} \mid \langle f \rangle < i(x)\}$ provide a Dedekind cut for $\boldsymbol{R}$. Thus there is a real number p such that if $x < p$, then $x \in A$, and if $p < x$, then $x \in B$. We claim that $\langle f \rangle - i(p)$ is infinitely small. Given $s > 0$, $\{n \in N \mid f(n) < s + p\} \in \mathscr{U}$ because $s + p \in B$; similarly $\{n \in N \mid p - s < f(n)\} \in \mathscr{U}$ because $p - s \in A$. Hence their intersection $\{n \in N \mid |f(n) - p| < s\} \in \mathscr{U}$, and $\langle f \rangle - i(p)$ is infinitesimal. Since X is closed, $p \in X$, and it follows that X is compact (Theorem 6).

THEOREM 9. *Every Cauchy sequence in $\boldsymbol{R}$ converges.*

Proof. Let $S: N \to \boldsymbol{R}$ be a Cauchy sequence. Then $\{S(n) \mid n \in N\}$ is bounded, as is well known, so there is a real number $x > 0$ such that $S(n) \in X = [-x, x]$ for each $n \in N$. We are going to use the compactness of X and Theorem 1 to prove this theorem. Let $\langle f \rangle = {}^*S(\langle \mu \rangle)$ where $\mu(n) = n$ for each $n \in N$: note that $\langle \mu \rangle \in {}^*N$ is infinitely large. Now $\langle f \rangle \in {}^*X$ since $\{n \in N \mid f(n) \in X\} = \{n \in N \mid S(\mu(n)) \in X\}$

$=N\in\mathscr{U}$. Since X is compact there is a unique $y\in X$ such that $\langle f\rangle - i(y)$ is infinitesimal, and we claim S converges to y. First note that for any infinitely large $\langle v\rangle\in{}^*N$ we have that ${}^*S(\langle v\rangle) - {}^*S(\langle\mu\rangle)$ is infinitesimal. For let $s>0$ be given, let $t\in N$ be such that $n,m\geqq t$ implies $|S(n)-S(m)|<s$; such t exists because S is a Cauchy sequence. Then

$$\begin{aligned}&\{n\in N \mid |{}^*S(\langle v\rangle)(n) - {}^*S(\langle\mu\rangle)(n)|<s\}\\ =\;&\{n\in N \mid v(n)\in N \text{ and } |S(v(n))-S(n)|<s\}\\ \supseteq\;&\{n\in N \mid v(n)\in N \text{ and } v(n),\, n\geqq t\}\\ =\;&\{n\in N \mid v(n)\in N\}\cap\{n\in N \mid v(n)\geqq t\}\cap\{t,t+1,\cdots\}\in\mathscr{U}.\end{aligned}$$

Thus for any infinitely large $\langle v\rangle\in{}^*N$, ${}^*S(\langle v\rangle) - {}^*S(\langle\mu\rangle)$ is infinitesimal, as is ${}^*S(\langle\mu\rangle) - i(y)$. Hence their sum ${}^*S(\langle v\rangle) - {}^*S(\langle\mu\rangle) + {}^*S(\langle\mu\rangle) - i(y) = {}^*S(\langle v\rangle) - i(y)$ is infinitesimal, and S converges to y (Theorem 1).

References

1. Edmund Landau, Foundations of Analysis, Chelsea, New York, 1960.
2. Abraham Robinson, Non-Standard Analysis, North-Holland, Amsterdam, 1966.

CORRECTION TO "FABER POLYNOMIALS AND THE FABER SERIES"

(This MONTHLY, 78 (1971) 577–596.)

J. H. CURTISS

Professor J. S. Frame has very kindly pointed out to me that in setting up the recurrence relation (2.4), (2.5) for the example of the three-cusped hypocycloid on page 583, I seem to have lost a coefficient 1/2. The recurrence should be $p_{n+3} = tp_{n+2} - (1/2)p_n$ and the first seven Faber polynomials should be

$$\begin{aligned}&p_1(2t) = t,\; p_2(2t) = t^2,\; p_3(2t) = t^3 - (3/2),\\ &p_4(2t) = t^4 - 2t,\; p_5(2t) = t^5 - (5/2)t^2,\\ &p_6(2t) = t^6 - 3t^3 + (3/4),\; p_7(2t) = t^7 - (7/2)t^4 + (7/4)t.\end{aligned}$$

Dr. Frame called attention to the fact that the general formula for $p_n(2t)$ is obtainable from his paper *Power Series Expansions for Inverse Functions* (this MONTHLY, 64 (1957) 236–240). It is

$$p_n(2t) = t^n + \sum_{k=1}^{[n/3]} (-1/2)^k n/k \binom{n-1-2k}{k-1} t^{n-3k}.$$

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306; notes are usually limited to three printed pages.

THE EXISTENCE OF FREE GROUPS

MICHAEL BARR, McGill University

Let X be a set. A free group generated by X consists of a group F and a function $f: X \to F$ satisfying the following condition:

For any group K and function $k: X \to K$, there is a unique group homomorphism $\phi: F \to K$ for which $\phi \circ f = k$.

It is a well-known theorem that *given any set X there is always a free group generated by X*. (It is an easy and instructive exercise for the reader to use the above definition to show that up to isomorphism, there is only one; also, the set $f(X)$ generates F, as explained below.) The usual proofs are by the construction of a semigroup of "words," multiplied by juxtaposition, that become a group modulo a rather complicated equivalence relation. Here we present a proof which never leads us out of the category of groups.

The proof is modeled after that of the general adjoint functor theorem of category theory and, as such, is readily adapted to solving any universal mapping problem in the category of groups, such as the existence of free products. It also works in any category consisting of all the algebras and algebra homomorphisms of any algebraic theory. (However, one must not foolishly exclude the empty set from such a category if it otherwise satisfies its axioms, for then the empty set might not generate a free algebra.) Thus included are all such categories as sets, sets with a basepoint (and base-point preserving functions), groups, abelian groups, rings, commutative rings, Lie rings, Jordan rings, algebras of these types, etc., each considered as a category with the evident definition of homomorphism.

1. Preliminaries. Let Γ be an index set, and let $\{G_\alpha\}$ be a family of groups indexed by Γ. The **product**

$$G = \prod_{\alpha \in \Gamma} G_\alpha$$

is the usual cartesian product with coordinate-wise multiplication. Each element of G is a Γ-tuple $\{x_\alpha\}$, where $x_\alpha \in G_\alpha$ for all $\alpha \in \Gamma$ and

$$\{x_\alpha\}\{y_\alpha\} = \{x_\alpha y_\alpha\}.$$

It is easily checked that G with this operation is a group. Also, each projection π_β

defined by

$$\pi_\beta\{x_\alpha\} = x_\beta$$

is a group homomorphism, $\pi_\beta \colon G \to G_\beta$.

Let G be a group and A a sub-*set* of G. We say that A **generates** G if no proper sub-*group* of G contains A.

If $g \colon X \to G$ is a map, we say that g **generates** G if the image $g(X) = \{a \in G \mid a = g(x)$ for some $x \in X\}$ generates G.

Suppose $g \colon X \to G$ and g does not necessarily generate G. Then g does, in a sense, generate a subgroup of G. Precisely:

PROPOSITION 1. *Let $g \colon X \to G$ be a map of a set X into a group G. Then there is a subgroup H of G and a map $h \colon X \to H$ such that h* **generates** *H and $g = j \circ h$, where j is the inclusion map.*

Proof. Let H be the intersection of all subgroups of G which contain $g(X)$. Clearly no proper subgroup of H contains $g(X)$, so $g(X)$ generates H. The rest is clear.

In this situation, we call H the subgroup of G **generated** by $g \colon X \to G$.

PROPOSITION 2. *Let X be a set. Then there exists a collection C of pairs (G_α, g_α), indexed by some set Γ such that*

(1) *each G is a group and $g_\alpha \colon X \to G_\alpha$ generates G_α;*

(2) *If K is any group and $k \colon X \to K$ generates K, then for some α there is an isomorphism ψ on G_α onto K such that $\psi \circ g_\alpha = k$.*

It is tempting to say "but this is obvious; simply take the collection of *all* pairs (G, g), where G is a group and $g \colon X \to G$ generates G." The sticky point is "all," which leads into the usual logical paradoxes. The proof of Proposition 2 will be postponed until the last section.

2. The proof. We state precisely the main result of this paper.

THEOREM. *Let X be a set. Then there exists a free group generated by X.*

Proof. Take a collection C given by Proposition 2. Form

$$G = \prod_{\alpha \in \Gamma} G_\alpha \text{ and } g = \prod_{\alpha \in \Gamma} g_\alpha.$$

Then G is a group and $g \colon X \to G$. Note that

$$\pi_\alpha \circ g = g_\alpha.$$

Let F be the subgroup of G generated by g. By Proposition 1, there is a map $f \colon X \to F$ such that f generates F and $g = j \circ f$, where j is the inclusion map of F in G. Note that

$$\pi_\alpha \circ j \circ f = \pi_\alpha \circ g = g.$$

Now let $k: X \to K$, where K is a group. We must prove that a unique homomorphism $\phi: F \to K$ exists such that $\phi \circ f = k$.

First we prove uniqueness. Suppose that we have two homomorphisms $\phi_i: F \to K$, where $\phi_i \circ f = k$ for $i = 1,2$. Let F_0 be the set of x in F such that $\phi_1(x) = \phi_2(x)$. Obviously F_0 is a subgroup of F. Since $\phi_1 \circ f = \phi_2 \circ f$, we have $f(X) \subseteq F_0$. But $f(X)$ generates F, hence $F_0 = F$, so $\phi_1 = \phi_2$. We pass to the existence proof.

First assume k generates K. Then Proposition 2 gives us an α and an isomorphism ψ on G_α onto K such that $k = \psi \circ g_\alpha$. Consider

$$F \xrightarrow[j]{} G = \prod G_\alpha \xrightarrow[\pi_\alpha]{} G_\alpha \xrightarrow[\Psi]{} K.$$

Let ϕ be the composite map, $\phi = \psi \circ \pi_\alpha \circ j$. Then ϕ is a homomorphism, and

$$\phi \circ f = \psi \circ \pi_\alpha \circ j \circ f = \psi \circ g_\alpha = k.$$

Thus such an f exists in this case.

Now suppose k does not generate K. By Proposition 1, there is a subgroup K' of K and a map $k': X \to K'$ which generates K' such that $j' \circ k' = k$, where $j': K' \to K$ is the inclusion map.

We know there is a homomorphism $\phi': F \to K'$ such that $\phi' \circ f = k'$. We set $\phi = j' \circ \phi'$; then $\phi: F \to K$ is a homomorphism, and

$$\phi \circ f = j' \circ \phi' \circ f = j' \circ k' = k.$$

3. The construction of C. In this section we wind up matters by proving Proposition 2. For any set S, let $|S|$ denote the cardinality of S.

LEMMA 1. *Let $k: X \to K$ generate K. Then $|K| \leqq \max(|X|, \aleph_0)$.*

Proof. Let $A = k(X)$. Then $|A| \leqq |X|$, so it is sufficient to show that $|K| \leqq \max(|A|, \aleph_0)$, where A is a set of generators of K. (Of course, if A is infinite, this is equivalent to the assertion that $|K| = |A|$.)

Let $A^{-1} = \{a^{-1} \mid a \in A\}$ and for B and C subsets of K, define $BC = \{bc \mid b \in B, c \in C\}$. Now let $A_1 = A \cup \{e\} \cup A^{-1}$, where e is the identity of K. Then define $A_2 = A_1A_1$, $A_3 = A_1A_2, \cdots, A_{n+1} = AA_n, \cdots$. Finally let $\bar{A} = \bigcup_{n=1}^{\infty} A_n$. By an easy induction, one sees that $A_n \subseteq A_n^{-1}$ and that $A_nA_m \subseteq A_{n+m}$. Thus $x, y \in \bar{A}$ implies $x^{-1} \in \bar{A}$ and $xy \in \bar{A}$, hence $\bar{A}$ is a subgroup of K. Since $A \subseteq \bar{A}$, this implies that $\bar{A} = K$. For any subsets B and C of K, the set BC is the image under the multiplication of $B \times C$, which implies that $|BC| \leqq |B| \cdot |C|$.

First we suppose that A is infinite. Then $|A_1| \leqq 1 + |A| + |A^{-1}|$, while $|A^{-1}| = |A|$, which gives $|A| = |A_1|$. Next $|A_2| \leqq |A_1| \cdot |A_1| = |A_1|$, while $A_1 \subseteq A_2$ (since $e \in A_1$, and $A_1 = eA_1 \subseteq A_1A_1$) implies $|A_1| \leqq |A_2|$, and then $|A_2| = |A|$. By induction, $|A_n| = |A|$ for all n, and thus $|\bar{A}| \leqq \Sigma |A_n| \leqq \aleph_0 \cdot |A| = |A|$.

When A is finite, the same argument shows instead that each A_n is finite, and the countable union $\bar{A}$ is at most countable.

LEMMA 2. *Let Y be any set. Then the collection of groups G whose underlying set is Y has cardinality at most*

$$c = |Y|^{|Y|^2}.$$

For a group is determined by its multiplication table, a function on $Y \times Y$ to Y. But c is the cardinality of the set of all such functions.

Proof of Proposition 2. For each cardinal s with $s \leqq \max(|X|, \aleph_0)$, choose a set Y_s with $|Y_s| = s$. Consider the set of all groups G such that the underlying set of G is some Y_s; by Lemma 2, this set exists. Then consider the collection C of all pairs (G_α, g_α), where G_α is one of these groups G_α and $g_\alpha: X \to G_\alpha$ generates G_α. The (2) of Proposition 2 is an immediate consequence of this construction and Lemma 1. The proof is complete.

INTEGERS WITH GIVEN INITIAL DIGITS

R. S. BIRD, Institute of Computer Science, London, England

Consider the following situation. Two mathematicians called X and Y are talking, and X announces that he has just computed a large prime, which he begins to recite to Y digit by digit. There are two possible responses open to Y. He can either wait until X has finished and then check the assertion, or he can interrupt X at some point in the recitation with the information that no prime can begin with those digits. The problems we are interested in are these:

(1) Assuming that Y knows his primes, can we prove that there is no sequence of digits that allows Y to interrupt X?

(2) Can the same be said about other sets of integers such as the squares, the factorial numbers, or the powers of 2?

To make things precise, suppose that S is an (infinite) set of positive integers. We shall say that S is **extendable in base b** if for each integer $x \geqq 1$, there are integers y and n, with $y < b^n$, such that $x b^n + y$ is in S.

If S is extendable in base b and consists of the integers $s_0, s_1, \cdots$, then

(1) for each integer $x \geqq 1$, there are integers m and n such that $b^n x \leqq s_m \leqq b^n(x+1)$.

Conversely, (1) implies that S is extendable in base b, as we can take y to be $s_m - b^n x$.

If we use the prime number theorem, the proof of the extendability of P, the set of primes, in every base is fairly easy.

THEOREM 1. *Let $\pi_S(n)$ be the number of members of S less than n. A sufficient*

condition for S to be extendable in every base is that if $\pi_S(n)/\pi_S(\lambda n) \to \theta(\lambda)$ *for all real* λ *satisfying* $1 \leqq \lambda \leqq 2$, *then* $\theta(\lambda) = 1$ *only if* $\lambda = 1$.

Proof. Assume that S is not extendable in some base b. Then by (1), there exists an x such that $\pi_S(b^n(x+1)) = \pi_S(b^n x)$ for all n. Let $\lambda_0 = (x+1)/x$ (whence $1 < \lambda \leqq 2$), and $m_n = b^n x$, so that

$$\pi_S(m_n)/\pi_S(\lambda_0 m_n) = 1 \text{ for all } m_n.$$

It follows that if $\pi_S(n)/\pi_S(\lambda n) \to \theta(\lambda)$, then $\theta(\lambda_0) = 1$ for some $\lambda_0 \neq 1$, which contradicts the hypothesis of the theorem.

Now the prime number theorem asserts that $\pi_P(n) \sim n/\log n$, whence $\pi_P(n)/\pi_P(\lambda n) \to 1/\lambda$, which is 1 only if $\lambda = 1$. Therefore P is extendable in every base. A similar argument shows that for each k, the set of kth powers is extendable in every base. However, in other interesting cases the ratio $\pi_S(n)/\pi_S(\lambda n)$ fails to converge, or converges to 1, for all λ, and a sharper condition is needed. The following theorem effectively characterises the extendable sets of numbers and reduces the question to a problem of Diophantine Approximation.

THEOREM 2. *A necessary and sufficient condition for the set* $S = \{s_0, s_1, \cdots\}$ *of positive integers to be extendable in base b is that the set of fractional parts of the real numbers* $\log_b s_0$, $\log_b s_1, \cdots$ *be dense in the unit interval.*

Proof. In the following, all logarithms are taken to the base b.

(a) *Necessity.* Suppose S is extendable in base b, so that condition (1) holds. Take logarithms and write $u_m = \log s_m$, $\alpha = \log x$, and $\delta(\alpha) = \log(1 + 1/x)$. Then, by assumption, for each α of the form $\log x$, there exist integers m and n such that

$$0 \leqq u_m - n - \alpha < \delta(\alpha). \tag{2}$$

If we write (z) for the fractional part and $[z]$ for the integral part of the real number z, then (2) can be expanded to

$$(\alpha) - (u_m) \leqq [u_m] - [\alpha] - n < \delta(\alpha) + (\alpha) - (u_m) \leqq \delta(\alpha) + (\alpha). \tag{3}$$

Since $(\alpha) - (u_m) > -1$, and $\delta(\alpha) + (\alpha) = \log(x+1) - [\log x]$, which has a maximum value of 1 (obtained when x is of the form $b^k - 1$, for some k), the above inequalities imply that $[u_m] - [\alpha] = n$, whence (2) can be simplified to

$$0 \leqq (u_m) - (\alpha) < \delta(\alpha). \tag{4}$$

Let ε be any positive real number, and x any integer. Define $\alpha_k = \log b^k x$. Clearly $(\alpha_k) = (\log x)$, for each k. Let n be any integer such that

$$\delta(\alpha_n) = \log(1 + 1/b^n x) < \varepsilon.$$

For such an n, there exists, by (4), an m such that $0 \leqq (u_m) - (\alpha_n) < \delta(\alpha_n)$; i.e., an m satisfying

$$(5) \qquad 0 \leqq (u_m) - (\log x) < \varepsilon.$$

Given ε, we can also find an integer x_0 such that $0 \neq (\log x_0) < \varepsilon$. The sequence of points

$$(\log x_0),\ (2\log x_0),\ (3\log x_0), \cdots$$

therefore marks a chain across the interval $(0,1)$, where the distance between consecutive points is less than ε. Hence, given any θ in $(0,1)$, one can find a number $x = x_0^k$, for some k, such that $0 \leqq \theta - (\log x) < \varepsilon$. It follows, using (5), that a number m exists such that

$$(6) \qquad \left|\theta - (u_m)\right| < \varepsilon.$$

Since θ and ε were arbitrary, (6) is just the condition for the set of fractional parts of $\log s_0, \log s_1, \cdots$ to be dense in the unit interval.

(b) *Sufficiency.* Suppose that (6) holds for arbitrary θ and ε. Let x be any positive integer, and take $\theta = (\log x)$. Then there must be an infinite number of integers m such that

$$0 \leqq (u_m) - (\log x) < \varepsilon,$$

for otherwise, we can construct an interval to the right of $(\log x)$ that contains no point of the form (u_m), contrary to assumption. If we take an $\varepsilon < \delta(\alpha)$ and an m such that $u_m \geqq \alpha$, where $\alpha = \log x$, then $n = [u_m] - [\alpha]$ is a non-negative integer that satisfies condition (1). Hence S is extendable in base b.

The following lemma is based on a proof by J. W. S. Cassels [**1**].

LEMMA. *Let U be a sequence $u_0, u_1, \cdots$ of real numbers of increasing size. A sufficient condition for the fractional parts of U to be dense in the unit interval is given by either*

(i) $\Delta u_n \to \theta$, *where θ is either irrational or zero, or*

(ii) $\Delta u_n \to \infty$, *and* $\Delta^2 u_n \to 0$. (*By definition,* $\Delta u_n = u_{n+1} - u_n$.)

Proof. To begin with, assume that $\Delta u_n \to 0$, and let ϕ be an arbitrary real number. Since $u_n \to \infty$, it is easy to verify that, given any $\varepsilon > 0$ and any integer m, there exist integers p and n_0 such that

$$(7) \qquad \left|u_n - \phi - p\right| < \varepsilon \text{ for all } n \text{ satisfying } n_0 \leqq n \leqq n_0 + m.$$

In particular, it follows that the fractional parts of U are dense in the unit interval. Actually (7) asserts slightly more, and this is used below.

In the case $\Delta u_n \to \infty$ and $\Delta^2 u_n \to 0$, it follows from (7) (with u_n replaced by Δu_n) that, given $\varepsilon > 0$ and m, there exist integers p and n_0 such that

$$\left|\Delta u_n - \phi - p\right| < \varepsilon/m \text{ for all } n \text{ satisfying } n_0 \leqq n \leqq n_0 + m.$$

The above statement (with $p = 0$ and $\phi = \theta$) also holds true in the third case $\Delta u_n \to \theta$, so that in either case, given ε and m, there exist p and n_0, and some irrational θ, such that

$$(8) \qquad \left| u_{n_0+k} - u_{n_0} - k\theta - kp \right| \leqq \sum_{r=0}^{k-1} \left| \Delta u_{n_0+r} - \theta - p \right| < \varepsilon$$

provided that $0 \leqq k \leqq m$.

Next, one version of Kronecker's theorem asserts that if θ is irrational, then given $\varepsilon > 0$, there is an n_1, such that for any real α there exist integers q and k_0, with $0 \leqq k_0 < n_1$, such that $\left| k_0\theta - \alpha - q \right| < \varepsilon$.

In substance, this says that the set of points $\{(\theta), (2\theta), \cdots\}$ is dense in the unit interval. For a proof see Cassels [2], or Hardy [3].

Now, if in (8) we take $m = n_1$, let β be arbitrary, and set $\alpha = \beta - u_{n_0}$, then Kronecker's theorem asserts the existence of integers q and k_0, with $k_0 \leqq n_1$, such that

$$\left| k_0\theta - \beta + u_{n_0} - q \right| < \varepsilon .$$

Setting $s = k_0 p + q$, it follows that

$$\left| u_{n_0+k_0} - \beta - s \right| \leqq \left| u_{n_0+k_0} - u_{n_0} - k_0 p - k_0\theta \right| + \left| k_0\theta - \beta + u_{n_0} - q \right| < 2\varepsilon .$$

Since β and ε were arbitrary and s is an integer, the lemma is proved.

THEOREM 3. *A sufficient condition for the set $S = \{s_0, s_1, \cdots\}$ of positive integers to be extendable in base b is that*

either (i) $s_{n+1}/s_n \to \theta$, *where $\theta = 1$ or θ is not a rational power of b,*

or (ii) $s_{n+1}/s_n \to \infty$ *and* $s_n s_{n+2}/s_{n+1}^2 \to 1$.

The proof is a straightforward consequence of the lemma and Theorem 2. The second condition is independent of b, and so asserts the extendability of S in every base.

Now we can show, for example, that the set of powers of a given integer p is extendable in base b, provided that p is not a power of b, as the first condition is satisfied. Also, the set of factorial numbers is extendable in every base as $(n+1)!/n! = n+1 \to \infty$ and $n!(n+2)!/(n+1)!^2 = (n+2)/(n+1) \to 1$.

References

1. J. W. S. Cassels, Personal Communication.
2. J. W. S. Cassels, An Introduction to Diophantine Approximation, Cambridge Tracts in Math., No. 45, C. U. P., England, 1965.
3. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th Edition, Clarendon Press, Oxford, England, 1960.

TORSION AT AN INFLECTION POINT OF A SPACE CURVE

RICHARD A. HORD, Langley Research Center

1. Introduction. In the differential geometry of Euclidean 3-space, the formula generally established for the torsion of a space curve (see, e.g., [1], page 23) is not applicable at a point where the curvature vanishes. Since some confusion has arisen (e.g., regarding conditions for the osculating plane to be stationary) because this special case is not considered in the standard works on differential geometry, a brief analysis of it is presented hereinafter. (See, however, the related and interesting discussion of singular points in [2], section 20, pages 41–43.) For a regular analytic space curve ([2], page 18) which is not a straight line, the existence of the torsion, for the special case in question, is demonstrated by proving that the direction of the binormal line is a continuous function of the arc length and that the torsion on the left and the torsion on the right exist and are equal; since this value of the torsion is equal to the limit of the complex torsion function at the point, the theory of functions implies straightforwardly that the zero-curvature point is a point of analyticity of the torsion function (cf. [2], page 41, where, perhaps for simplicity, torsion on the left and right were not considered). The formula subsequently derived for the torsion is applicable at a point of zero curvature; at an ordinary point, it reduces to the familiar expression. As a physical example, a point of zero curvature exists in the trajectory of a particle whose velocity vector is instantaneously parallel to the resultant of all forces acting on the particle.

2. Preliminary analysis. The position vector of a point on a regular analytic space curve Γ which is not a straight line has (locally, at least) a convergent Maclaurin representation of the form

$$r = r_0 + r_0's + r_0^{(k)}\frac{s^k}{k!} + r_0^{(k+1)}\frac{s^{k+1}}{(k+1)!} + \cdots \quad (k \geqq 2), \tag{2.1}$$

where the arc length s has domain consisting (at least) of an interval including zero, and where the vector $r_0^{(k)}$ is not zero. Primes and superscripts in parentheses signify derivatives with respect to s. If $k = 2$, r_0 is called an **ordinary point.** If $k > 2$, so that r_0'' and the curvature at r_0 are zero, then r_0 is said to be a **singular point** [2] or an **inflection point** [3].

In [1] and [2], for example, the curvature is defined in such a way that it is nonnegative and the principal normal vector n and the binormal vector b may reverse their senses at an inflection point. In [3], the curvature can change sign to allow the aforementioned unit vectors to be continuous functions of s for a regular analytic curve. The principal normal and binormal *lines* are not subject to ambiguity. The **principal normal** is the line of the vector

$$r'' = r_0^{(k)}\frac{s^{k-2}}{(k-2)!} + r_0^{(k+1)}\frac{s^{k-1}}{(k-1)!} + \cdots \quad (k \geqq 2), \tag{2.2}$$

which is an analytic (i. e., all components analytic) function of s and vanishes at most at isolated values of s; at an inflection point, say $s = 0$ with $k > 2$, the principal normal is the line of the vector $r_0^{(k)}$ which is non-zero and normal to the unit tangent vector at the point, that is, r_0' [3]. Thus the osculating plane exists at every value of s [3]. The binormal line has the direction of $r' \times r''$, which is an analytic function of s and vanishes at most at isolated values of s; at an inflection point, say $s = 0$ with $k > 2$, the binormal line properly has the direction of $r_0' \times r_0^{(k)}$, which is non-zero. Thus the rectifying plane exists at every value of s. Now r and its derivatives are analytic functions of s. Since, by equation (2.2),

$$\frac{r''}{|r''|} = \left(\frac{s}{|s|}\right)^{k-2} \frac{r_0^{(k)} + (s/(k-1))r_0^{(k+1)} + \cdots}{\left|r_0^{(k)} + (s/(k-1))r_0^{(k+1)} + \cdots\right|} \qquad (s \neq 0) \tag{2.3}$$

it is clear that the directions of the principal normal and binormal *lines* are continuous functions of s; moreover, the continuity of the principal normal and binormal *vectors* is assured at points where k is an even integer, for either of the ways mentioned for defining the curvature. If the curvature is defined so as to be non-negative, then n has the sense of r'' and it follows from equation (2.3) that n reverses across $s = 0$ when k is odd; similarly, b reverses in such a case. In any case, n has a limit on the left, say n_-, and a limit on the right, say n_+, at $s = 0$; similarly, b has left and right hand limits, say b_- and b_+.

3. Torsion. At an ordinary point of a regular space curve (of class C^3), the torsion, by definition, satisfies the vector equation

$$\frac{db}{ds} = -\tau n \qquad \left(b = \frac{r' \times r''}{|r' \times r''|}\right) \tag{3.1}$$

and can be expressed in the familiar form

$$\tau = \frac{[r', r'', r''']}{r'' \cdot r''}, \tag{3.2}$$

where the brackets denote the triple scalar product.

(Usually the sense of b is chosen to make the moving trihedron always right-handed (or, perhaps, always left-handed), so that τ may be of either sign (or zero). However, having chosen the unit tangent and principal normal vectors, the sign of the torsion could be kept non-negative by replacing b by $-b$ as required; both right- and left-handed moving trihedra would thereby be admitted.)

At an arbitrary point, say $s = 0$, of a regular analytic space curve, equations (2.1) and (3.2) yield

$$\tau_- = \tau_+ = \lim_{s \to 0} \tau = \frac{1}{k-1} \frac{[r_0', r_0^{(k)}, r_0^{(k+1)}]}{r_0^{(k)} \cdot r_0^{(k)}}, \tag{3.3}$$

where τ_- and τ_+ denote the limits of τ on the left and on the right, respectively, at $s = 0$. Let τ_l and τ_r denote the torsion on the left and on the right, respectively; then

$$(3.4) \quad -\tau_l \boldsymbol{n}_- = \lim_{s\to 0-} \frac{\boldsymbol{b}_- - \boldsymbol{b}}{0 - s} = \lim_{s\to 0-} \frac{d\boldsymbol{b}/ds}{ds/ds} = \lim_{s\to 0-} \frac{d\boldsymbol{b}}{ds} = -(\tau\boldsymbol{n})_- = -\tau_-\boldsymbol{n}_-$$

and

$$(3.5) \quad -\tau_r \boldsymbol{n}_+ = \lim_{s\to 0+} \frac{\boldsymbol{b} - \boldsymbol{b}_+}{s - 0} = \lim_{s\to 0+} \frac{d\boldsymbol{b}/ds}{ds/ds} = \lim_{s\to 0+} \frac{d\boldsymbol{b}}{ds} = -(\tau\boldsymbol{n})_+ = -\tau_+\boldsymbol{n}_+ .$$

(Equation (3.3) can also be obtained from equation (3.2) by a well-known generalization of L'Hospital's rule. Symbolically, equation (3.2) has the form $\tau = [1,2,3]/(2\cdot 2)$. Now

$$\left\{\frac{d^{2(k-2)}}{ds^{2(k-2)}}[1,2,3]\right\}_0 = \frac{[2(k-2)]!}{(k-2)!(k-2)!}[1,k,k+1]_0 + \frac{[2(k-2)]!}{(k-1)!(k-3)!}[1,k+1,k]_0$$

$$\left\{\frac{d^{2(k-2)}}{ds^{2(k-2)}}(2\cdot 2)\right\}_0 = \frac{[2(k-2)]!}{(k-2)!(k-2)!}(k\cdot k)_0 \neq 0,$$

where the subscripts 0 denote values at $s = 0$. Equation (3.3) follows.)

Since $\boldsymbol{n}_-$ and $\boldsymbol{n}_+$ are unit vectors (and not zero vectors), it follows that

$$(3.6) \qquad \tau_l = \tau_-, \ \tau_r = \tau_+ .$$

Equations (3.3) and (3.6) complete the proof of the following:

THEOREM. *At a given point, say $s = 0$, of a regular analytic space curve which is not a straight line, the torsion on the left, τ_l, the torsion on the right, τ_r, and* $\lim_{s\to 0}\tau$, *the limit of the torsion function τ at $s = 0$, all exist and are equal to*

$$(3.7) \qquad \tau_0 = \frac{1}{k-1}\frac{[\boldsymbol{r}'_0, \boldsymbol{r}_0^{(k)}, \boldsymbol{r}_0^{(k+1)}]}{\boldsymbol{r}_0^{(k)}\cdot \boldsymbol{r}_0^{(k)}},$$

which is properly termed the **torsion** *of the curve at $s = 0$; the notation is that of equation* (2.1). *Furthermore, the torsion is an analytic function of the arc length.*

The last statement follows from the observation that, even if $s = 0$ is a point at which $\boldsymbol{r}'' = 0$ ($k > 2$), the singularity in equation (3.2) is, from the complex variable standpoint, both isolated and removable.

If $\boldsymbol{r}''_0 \neq 0$ ($k = 2$), equation (3.7) has the form of equation (3.2), the expression for the torsion at an ordinary point.

COROLLARY. *The osculating plane of a regular analytic space curve which is not a straight line is stationary at a given point, say $s = 0$, if and only if*

$$[r'_0, r_0^{(k)}, r_0^{(k+1)}] = 0, \tag{3.8}$$

that is, $r_0^{(k+1)}$ *is a linear combination of the linearly independent vectors* r'_0 *and* $r_0^{(k)}$. (*The notation is that of equation* (2.1).)

Equation (3.8) corresponds to the cases in [2], pages 42–43, for which the integer n exceeds unity.

If the assumption that the curve is analytic everywhere is relaxed at a single point and replaced by the assumption that the curve is of class C^∞, then the osculating plane need not exist at the point, the direction of the binormal line need not have a two-sided limit at the point, and the torsion, if it could reasonably be defined, might be of infinite magnitude at the point. These observations follow from consideration of the point (an inflection point), $u = 0$, of the curve γ:

$$r(u) = \begin{cases} (u, e^{-1/u^2}, 0), & u < 0 \\ (0,0,0), & u = 0 \\ (u, 0, e^{-1/u^2}), & u > 0 \end{cases} \tag{3.9}$$

which is discussed briefly in [3], pages 9–10.

The author wishes to thank Professor T. J. Willmore for his suggestions on the paper.

References

1. L. P. Eisenhart, An Introduction to Differential Geometry, Princeton University Press, 1940 (2nd ed., 1947).
2. W. C. Graustein, Differential Geometry, Macmillan, New York, 1935, (Dover ed., 1966).
3. T. J. Willmore, An Introduction to Differential Geometry, Oxford University Press, 1959.

ON WHITNEY'S LINE GRAPH THEOREM

R. L. HEMMINGER, Vanderbilt University

In 1932, Whitney [6] proved that, with just four exceptions, line isomorphisms between finite connected graphs are induced by (point) isomorphisms. His argument involves many cases and is rather long. Better proofs are now known. One, due to Krause [4], is presented in Ore's book [5], and a very short and elegant proof, due to Jung [3], is given in Harary's book [1]. Moreover, Jung's proof holds for infinite graphs.

We shall show here how the exceptional cases arise. They are usually handled (cf. [5] p. 246) as follows: "We leave it to the reader to verify that only in these [four] instances can such correspondences occur in graphs whose orders do not exceed four." Whitney does this in his original article but not in a way that clarifies why there are anomalies.

We shall follow the notation and terminology of Harary [1]. In particular, $S(v)$ denotes the set of lines of G that are incident with the point v. We shall call a set S of lines of G a star of G if $S \subseteq S(v)$ for at least one v of G, and we shall say that a function σ, from the set of lines of G into the set of lines of G', preserves stars if the set $\sigma(S)$ is a star whenever the set S is a star.

THEOREM 1. *Let σ be a one-to-one function from the set of lines of G onto the set of lines of G', where G and G' are connected graphs. Then σ is induced by an isomorphism of G onto G' if and only if σ and σ^{-1} preserve stars.*

Proof. The condition is clearly necessary so we assume that σ and σ^{-1} preserve stars. Hence, for each point v in G, there is at least one point v' in G' such that $\sigma(S(v)) \subseteq S(v')$. Moreover v' is uniquely determined by v if $\mathrm{dg}(v) > 1$, since $S(v') \cap S(v'')$ is a singleton set if $v' \neq v''$. Thus, when $\mathrm{dg}(v) > 1$, we have $\mathrm{dg}(v') \geqq \mathrm{dg}(v) > 1$, so we must have $\sigma^{-1}(S(v')) \subseteq S(v)$. We conclude that the function σ determines a unique function σ^* from the set of points of G that have degree greater than one onto the set of points of G' that have degree greater than one, such that $\sigma(S(v)) = S(\sigma^*(v))$ (σ^* is an onto function because σ^{-1} enjoys the same properties as σ).

We hereafter assume that $|V(G)| \geqq 3$, otherwise the result is trivial. Thus, if $x = uv$ is a line in G with $\mathrm{dg}(v) = 1$, then $\mathrm{dg}(u) > 1$, so $\sigma(x) \in \sigma(S(u)) = S(\sigma^*(u))$. By the results of the last paragraph, we must have $\sigma(x) = u'v'$ where $u' = \sigma^*(u)$ and $\mathrm{dg}(v') = 1$. Therefore, if we extend σ^* by defining $\sigma^*(v) = v'$, then we conclude that σ determines a unique function, which we still denote by σ^*, from the points of G into the points of G' such that $\sigma(S(v)) = S(\sigma^*(v))$. However, since $|V(G)| \geqq 3$, $S(u) = S(v)$ if and only if $u = v$. Thus σ^* is a one-to-one function, and hence an onto function, from the points of G onto the points of G'.

It is now obvious that σ^* is an isomorphism of G onto G' that induces the function σ.

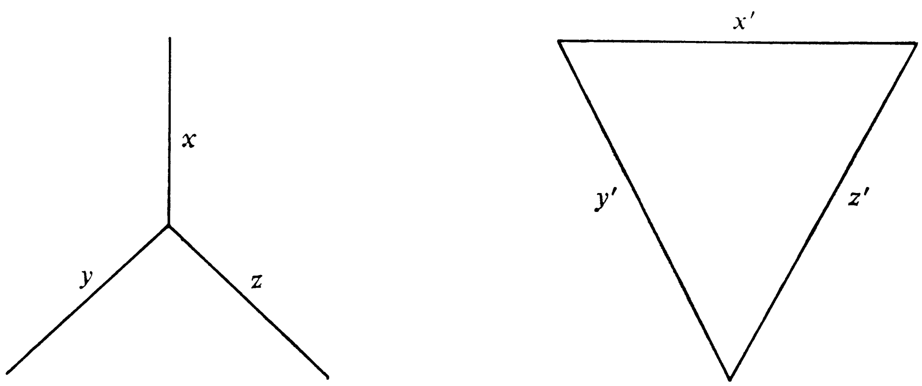


FIG. 1

This argument is essentially the one used by Jung [3] in a portion of his proof of Whitney's theorem. However, by using it to prove the theorem above we can now see how the exceptional cases arise in Whitney's theorem.

Observe first that a function σ as in Theorem 1 is a line isomorphism (i.e., lines x and y are adjacent in G if and only if the lines $\sigma(x)$ and $\sigma(y)$ are adjacent in G'). Thus the line isomorphisms that are not induced by isomorphisms are precisely those having the property that they, or their inverses, fail to preserve stars.

Suppose that σ is a line isomorphism and v is a point such that $\sigma(S(v))$ is not a star. Then $\mathrm{dg}(v) = 3$ and $\sigma(S(v))$ is the line set of a triangle: for $\sigma(S(v))$ is obviously a star if $\mathrm{dg}(v) = 1$ or 2, and the only way for 4 or more lines to be pairwise adjacent is in a star. Let $S(v) = \{x, y, z\}$.

If $S(v)$ is the line set of G, then we have the first and basic exceptional case. It is illustrated in Figure 1.

Suppose that this is not the case. Then there is a line w in G that is adjacent to one of the lines of the set $S(v)$ since G is connected. But w must then be adjacent to two of the elements of $S(v)$ since $\sigma(w)$ must be adjacent to exactly two of the elements of $\sigma(S(v))$. There are just three such lines possible in G. Figures 2, 3, and 4 illustrate the remaining exceptions to Whitney's theorem. They correspond, respectively, to the existence in G of one, two, or three such lines of this type.

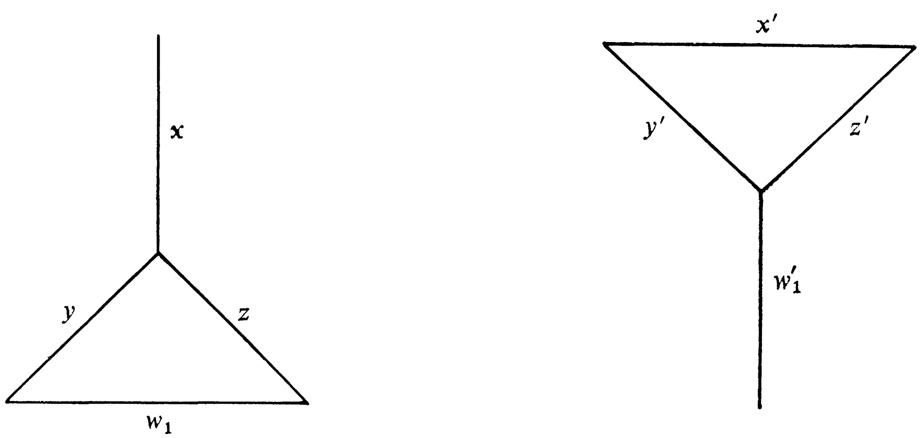


FIG. 2

Since the pairs in Figures 2, 3, and 4 are isomorphic we have the following result:

COROLLARY (Whitney's Theorem for Line Graphs). *If G and G' are connected graphs with isomorphic line graphs then G and G' are isomorphic graphs unless one is isomorphic to K_3 and the other isomorphic to $K_{1,3}$.*

With only a little more effort one can prove the following generalization of Theorem 1:

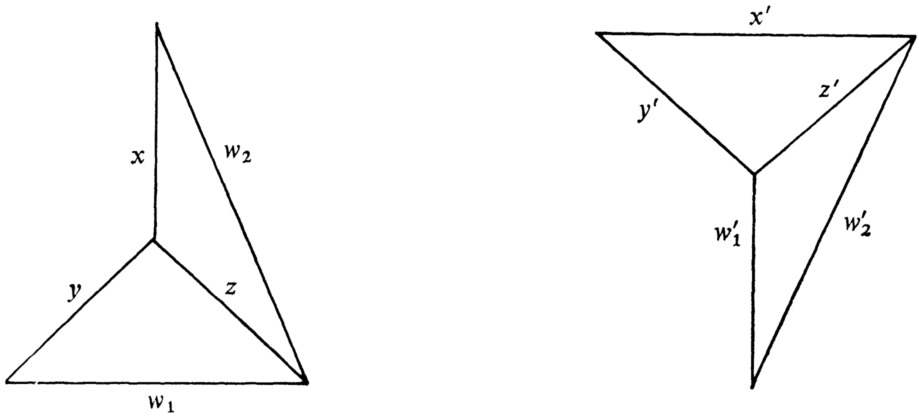


FIG. 3

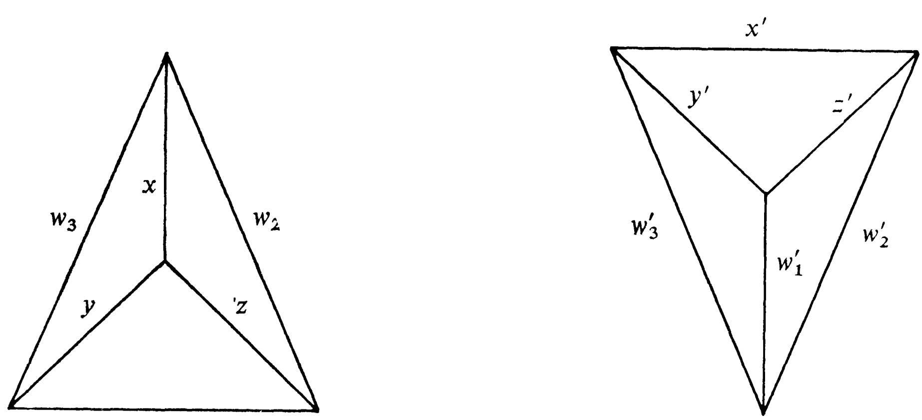


FIG. 4

THEOREM 2. *Let σ be a one-to-one function from the set of lines of G onto the set of lines of G′, where G and G′ are connected pseudographs (loops and multiple lines are allowed). Then σ is induced by an isomorphism of G onto G′ if and only if σ and σ^{-1} preserve loops, multiple lines, and stars.*

Using this result in the same way as Theorem 1 was used above, the author [2] has described the line isomorphisms between pseudographs that are not induced by isomorphisms. These can be classified into nine classes of pseudographs, each of which is closely related to one of the exceptional graphs in the Whitney-Jung theorem; however, some of the classes are infinite.

This work was done while the author was an NSF Science Faculty Fellow, visiting the University of California at Berkeley.

References

1. F. Harary, Graph Theory, Addison-Wesley, Reading, Mass., 1969.
2. R. Hemminger, Isomorphism-induced line isomorphisms on pseudographs, submitted to Czechoslovak Math. J.
3. H. Jung, Zu einem Isomorphiesatz von Whitney für Graphen, Math. Ann., 164 (1966) 270–271.
4. J. Krause, Démonstration nouvelle d'un théorème de Whitney sur les réseaux, Mat. Fiz. Lapok, 50 (1943) 75–85.
5. O. Ore, Theory of Graphs, Amer. Math. Soc. Colloq. Publ. 38, Providence, 1962.
6. H. Whitney, Congruent graphs and the connectivity of graphs, Amer. J. Math., 54 (1932) 150–168.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

A PACKING PROBLEM FOR TRIANGULAR MATRICES

W. KLOTZ AND L. LUCHT, Technische Universität Clausthal, Germany

Let $T(n)$ be the maximal number of ones in a triangular $n \times n$ matrix such that no rectangle is formed by ones. The following bounds of $T(n)$ are known.

$$(1) \qquad \tfrac{1}{2}n^{3/2} + o(n^{3/2}) \leqq T(n) \leqq \tfrac{2}{3}n^{3/2} + O(n^{5/4}).$$

It would be interesting to find an asymptotic formula for $T(n)$. As is seen from [4] this might improve some estimates for finite lattices consisting of a given number of elements. The problem of determining $T(n)$ belongs to the following, very general type of questions. Let S be a bounded subset of the real space R^m. What is the maximal number of lattice points in S such that a certain forbidden configuration is not formed?

Many papers dealing with matrix problems similar to ours originate in a problem posed by Zarankiewicz [5]. The arguments developed in these papers can be applied to certain extremal problems of graphs (see the papers of Brown [1] and Erdös, Rényi, Sós [2]). Guy [3] gives a survey and further references concerning this subject.

We outline the proof of (1). First we consider a rectangular $n \times m$ matrix. Denote by s_i the number of ones in the ith line. Suppose the matrix does not contain

a rectangle formed by ones. Then every two columns have at most one pair of ones both in the same line. This means

$$\sum_{i=1}^{n} \binom{s_i}{2} \leqq \binom{m}{2},$$

from which we prove

$$\sum_{i=1}^{n} s_i \leqq m\sqrt{n} + n.$$

In order to get an estimate for a triangular $n \times n$ matrix we cover the triangle by rectangles of size $k \times ik$, $i = 1, 2, \cdots, [n/k] + 1$ ($[x]$ denotes the integral part of x). If $T(n)$ is the total number of ones in this matrix, then

$$T(n) \leqq \sum_{i=1}^{[n/k]+1} (k\sqrt{ik} + ik).$$

Taking $k = [n^{3/4}]$ proves the upper bound in (1).

Assume that n is approximated by $y = q^2 + q + 1$ with q being a prime power. There is a projective plane of y points and y lines. The incidence matrix of this plane contains $(q+1)(q^2+q+1)$ ones. Since two lines meet in exactly one point there is no rectangle formed by ones. By considering one half of this incidence matrix the lower bound of $T(n)$ becomes evident.

References

1. W. G. Brown, On graphs that do not contain a Thomson graph, Canad. Math. Bull., 9 (1966) 281–285.

2. P. Erdös, A. Rényi, and V. T. Sós, On a problem of graph theory, Studia Sci. Math. Hungar., 1 (1966) 215–235.

3. R. K. Guy, A many-faceted problem of Zarankiewicz, Lecture Notes in Mathematics, 110 (1969) 129–148.

4. W. Klotz and L. Lucht, Endliche Verbände, J. Reine Angew. Math., 247 (1971) 58–68.

5. K. Zarankiewicz, Problem P101, Colloq. Math., 2 (1951) 301.

A PROBLEM IN GROUP THEORY

R. Hirshon, Polytechnic Institute of Brooklyn

A group G is said to be **hopfian** if endomorphism of G onto G is an automorphism. Equivalently, G is hopfian if $G/K \approx G$ implies $K = 1$; that is, G is not isomorphic to a proper factor group of itself.

We pose the following problem: If $G = H \times C_\infty$ is the direct product of a hopfian group H with an infinite cyclic group, is G hopfian? At first thought, one feels that since G is formed in such a simple manner from H, the answer to the question must be yes. On the other hand, it is possible to have (finitely generated) groups A and A_1

such that

$$A \times C_\infty \approx A_1 \times C_\infty$$

but $A \neq A_1$ [**17**]. If we could choose A and A_1 as above with A hopfian and A_1 a proper homomorphic image of A, then we would see that the answer to our question is no. In view of this consideration a conjecture that the answer is no is perhaps not too wild. Indeed some anomalous situations do come up with regard to hopficity. For example, A. L. S. Corner [**7**] has given an example of a hopfian abelian group A such that $A \times A$ is nonhopfian. Examples of finitely generated nonhopfian groups have been constructed by several mathematicians ([**6**], [**11**], [**19**]).

There are a few conditions which will guarantee a yes answer to our question. The simplest is that H be finitely generated [**16**] or that H be abelian [**13**].

References

1. Michael Anshel, Non hopfian groups with fully invariant kernels, Ph. D. dissertation, Adelphi University, 1967.

2. ———, The endomorphisms of certain one relator groups and the generalized hopfian problem, Bull. Amer. Math. Soc., 77(1971) 348–350.

3. R. Baer, Groups without proper isomorphic quotient groups, Bull. Amer. Math. Soc., 50 (1944) 267–278.

4. G. Baumslag, Hopficity and abelian groups, Topics in abelian groups, Scott-Foresman, Chicago, 1963, pp. 331–335.

5. ———, A non-hopfian group, Bull. Amer. Math. Soc., 73 (1967) 402–418.

6. G. Baumslag and D. Solitar, Some two generator one relator non-hopfian groups, Bull. Amer. Math. Soc., 68 (1962) 199–201.

7. A. L. S. Corner, Three examples on hopficity in torsion-free abelian groups, Acta Mathematica, 16 (1965) 303–310.

8. I. Dey, Free products of hopfian groups, Math. Zeitschr., 85 (1964) 274–284.

9. S. Dick, Dual-hopfian abelian groups, Ph. D. dissertation, Adelphi University, June 1968.

10. K. Frederick, The hopfian property for a class of fundamental groups, Comm. Pure Appl. Math., 16 (1963) 1–8.

11. G. Higman, A finitely related group with an isomorphic proper factor group, J. London Math. Soc., 22 (1951) 59–61.

12. R. Hirshon, Some results on direct products of hopfian groups, Ph. D. dissertation, Adelphi University, New York 1967.

13. ———, Some theorems on hopficity, Trans. Amer. Math. Soc., 141 (1969) 229–244.

14. ———, On hopfian groups, Pacific J. Math., 32 (1970) 753–766.

15. ———, The center and commutator subgroup in hopfian groups, Ark. Mat. (Stockholm), 9 (1971), 181–192.

16. ———, A conjecture on hopficity and related results, Arch. Math. (Basel), to appear.

17. ———, On cancellation in groups, this MONTHLY, 76 (1969) 1037–1039.

18. ———, Some new groups admitting essentially unique directly indecomposable decompositions, Math. Ann., to appear.

19. B. H. Neumann, A two generator group isomorphic to a proper factor group, J. London Math. Soc., 25 (1950) 247–248.

20. M. Orzech, Onto endomorphisms are isomorphisms, this MONTHLY, 78 (1971) 357–361.

21. M. Orzech and L. Riber, On residual finiteness and the hopfian property in rings, J. Algebra, 15 (1970) 81.

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306; notes are usually limited to three printed pages.

A VERSATILE VECTOR MEAN VALUE THEOREM

D. E. SANDERSON, Iowa State University

If a particle moves smoothly in n-space and at two points in time its velocity is orthogonal to a given direction, then so must its acceleration be at some intermediate time. The following easily proved extension of Rolle's theorem embodies this principle for arbitrary dimension and orders of differentiation (the one-dimensional case reduces to Rolle's theorem if orthogonality is interpreted as meaning the (inner) product of the vectors is zero). The two-dimensional version affords a simple way to present the elementary applications or forms of the usual mean value theorems.

THEOREM 1. *Suppose $\mathbf{v}\colon [a,b] \to R^n$ is a k times differentiable n-dimensional vector-valued function and $\mathbf{v}(a)$, $\mathbf{v}(b)$ and the first $k-1$ derivatives of $\mathbf{v}$ at a are orthogonal to a non-zero vector $\mathbf{v}_0$. Then for some c between a and b, $\mathbf{v}^{(k)}(c)$ is orthogonal to $\mathbf{v}_0$.*

Proof. Let $F(t) = \mathbf{v}(t) \cdot \mathbf{v}_0$ denote the inner (dot) product of the vectors $\mathbf{v}(t)$ and $\mathbf{v}_0$. Then, since the vanishing of $F^{(m)}(t) = \mathbf{v}^{(m)}(t) \cdot \mathbf{v}_0$ is equivalent to orthogonality of $\mathbf{v}^{(m)}(t)$ and $\mathbf{v}_0$, we have $F(b) = F(a) = F'(a) = \cdots = F^{(k-1)}(a) = 0$. Successive applications of Rolle's theorem give points $c_0 = b, c_1, \cdots, c_k = c$ such that $F^{(m)}(c_m) = 0$ and $a < c_m < c_{m-1}$ for $m = 1, \cdots, k$. Thus $\mathbf{v}^{(k)}(c)$ is orthogonal to $\mathbf{v}_0$ and the proof is complete.

To illustrate the ease with which standard mean value results can be obtained from this theorem (with $n = 2$) let us simplify the form by translating coordinates in the domain and range of $\mathbf{v}$ so that a is replaced by 0, b by $h = b - a$, and $\mathbf{v}(0)$ by the origin of R^2. If we write $\mathbf{v}(t) = (f(t), g(t))$ where $f(0) = g(0) = 0$, and assume $\mathbf{v}(h)$ is non-zero, then we may use $(g(h), -f(h))$ for $\mathbf{v}_0$ so that $F(t) = f(t)g(h) - g(t)f(h)$ and the orthogonality condition in the conclusion becomes $f(h)g^{(k)}(c) = g(h)f^{(k)}(c)$. This remains true, trivially, but of little use if $\mathbf{v}(h)$ is the zero vector.

Applications. (1) The ordinary mean value theorem for a function f, differentiable on $[0, h]$ (where $f(0) = 0$) is obtained by setting $k = 1$, $g(t) = t$: $f(h) = hf'(c)$.

(2) The Cauchy or generalized mean value theorem results from setting $k = 1$: $f(h)g'(c) = g(h)f'(c)$ (where $f(0) = g(0) = 0$).

(3) From (2) and appropriate conditions on f and g, one can of course write $f(h)/g(h) = f'(c)/g'(c)$ and derive L'Hospital's Rule.

For applications involving values of k greater than one (and $n = 2$, still) it should be observed that the condition on v and its first $k-1$ derivatives at a requires them to all be parallel. In particular, the theorem is applicable whenever the values of f and its first $k-1$ derivatives at a are equal to the respective values of g and its first $k-1$ derivatives at a. We state this as the next application, continuing to use the notationally simpler case, $a = 0$.

(4) If $f^{(m)}(0) = g^{(m)}(0)$ for $m = 0, 1, \cdots, k-1$ ($f^{(0)} = f$, etc.) and $f^{(k)}(t)$, $g^{(k)}(t)$ exist for $t \in [0, h]$, then $f(h)g^{(k)}(c) = g(h)f^{(k)}(c)$ for some c between 0 and h.

(5) Taylor's Formula for a k times differentiable function ϕ follows from (4) if we set $f(t) = \phi(t) - \sum_{s=0}^{k-1} \phi^{(s)}(0)t^s/s!$ and $g(t) = t^k$.

Proof. Since $f^{(m)}(t) = \phi^{(m)}(t) - \sum_{s=m}^{k-1} \phi^{(s)}(0)t^{s-m}/(s-m)!$, we have $f^{(m)}(0) = 0 = g^{(m)}(0)$ for $m = 0, 1, \cdots, k-1$ and (4) applies, giving $f(h)k! = h^k f^{(k)}(c) = h^k \phi^{(k)}(c)$, hence

$$\phi(h) = \sum_{s=0}^{k-1} \phi^{(s)}(0)h^s/s! + \phi^{(k)}(c)h^k/k!.$$

(6) The standard formula for the error in Simpson's Rule for approximating the integral of a four times differentiable function ϕ on the interval $[-h, h]$ follows from Corollary 4 by setting

$$f(t) = (t/3)[\phi(-t) + 4\phi(0) + \phi(t)] - \int_{-t}^{t} \phi \quad \text{and } g(t) = t^5.$$

Proof. Differentiating, one finds that f and its first three derivatives vanish at 0. In particular, $f'''(t) = [\phi'''(t) - \phi'''(-t)]t/3$. Applying (4) with $k = 3$ and using the mean value theorem (i.e., (1) modified to apply to the interval $[-c, c]$) gives

$$f(h) \cdot 60c^2 = [\phi'''(c) - \phi'''(-c)]h^5c/3 = 2c\phi^{(4)}(\bar{c})h^5c/3,$$

or

$$f(h) = h^5\phi^{(4)}(\bar{c})/90,$$

where $\bar{c} \in (-c, c) \subset (-h, h)$. This is the standard formula for the error $f(h)$ in Simpson's Rule.

Note that in the proof of (6) we could just as well apply the theorem with $k = 4$, and it would be more natural to do so. However, this leads to the more complicated form

$$f(h) = [2\phi^{(4)}(\bar{c}) + \phi^{(4)}(c) + \phi^{(4)}(-c)]h^5/360$$

and the same estimate $|f(h)| < Mh^5/90$, where M is the maximum of $|\phi^{(4)}(t)|$ for $-h < t < h$.

(7) The standard formula for the error in the Trapezoidal Rule for approximating the integral of a twice differentiable function ϕ on the interval $[-h, h]$

follows from (4) by setting

$$f(t) = [\phi(-t) + \phi(t)]/2 - \int_{-t}^{t} \phi \text{ and } g(t) = t^5 \text{ (and } k = 1).$$

The proof of (7) parallels that of (6), the error formula being $\frac{2}{3}h^3\phi''(\bar{c})$ for some $\bar{c} \in (-h, h)$. The corresponding formulas for an arbitrary interval divided into several (equal) subintervals are easily obtained if $\phi^{(4)}$ (respectively, ϕ'') is continuous on the interval (see problem 9 section 8.22 of [**1**]). The fact that the hypothesis of the theorem is satisfied for a higher value of k than is used in the proofs of (6) and (7) suggests that a sharper error estimate may be possible but the note preceding (7) does not bear this out.

Using Theorem 1 with $k = 1$ in much the same way that Rolle's theorem was used in proving Theorem 1, the following variation can be proved:

THEOREM 2. *Suppose* $\mathbf{v}$: $[a, b] \to R^n$ *is a* k *times differentiable* n*-dimensional vector-valued function which is orthogonal to a non-zero vector* $\mathbf{v}_0$ *at* $k+1$ *distinct points of* $[a, b]$. *Then for some* c *between* a *and* b, $\mathbf{v}^{(k)}(c)$ *is orthogonal to* $\mathbf{v}_0$.

Theorem 2 can be used to obtain the error formula for polynomial interpolation given in Theorem 8–3 of [**1**].

Reference

1. T. M. Apostol, Calculus, vol. 2, Blaisdell, New York, 1962.

A NOTE ON UNIFORM STRUCTURES OF TOPOLOGICAL GROUPS

J. S. YANG, University of South Carolina

We present here an extension of an exercise in [1, 4.24, page 28] which states that if there are sequences $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ in a T_0 topological group G such that $\lim_{n\to\infty} x_n y_n = e$ and $\lim_{n\to\infty} y_n x_n = z \neq e$, then the left and right uniform structures of G are inequivalent.

It is well known that a topological group G has equivalent left and right uniform structures if and only if for each neighborhood U of the identity e, there is a neighborhood V of e such that $xVx^{-1} \subset U$ for all $x \in G$ (cf. [1], 4.14, page 22).

THEOREM. *A topological group* G *has inequivalent left and right uniform structures if and only if there are nets* $\{x_\alpha\}$ *and* $\{y_\alpha\}$ *in* G *such that* $\{x_\alpha y_\alpha\}$ *converges to the identity* e *but* e *is not a cluster point of the net* $\{y_\alpha x_\alpha\}$.

Proof. Suppose there are nets $\{x_\alpha\}$ and $\{y_\alpha\}$ such that $\{x_\alpha y_\alpha\}$ converges to e, but e is not a cluster point of the net $\{y_\alpha x_\alpha\}$. Then there is a neighborhood U of e in G such that $\{y_\alpha x_\alpha\}$ is eventually in $W = G - U$. Let V be an arbitrary neighbor-

hood of e. If β is so chosen that $x_\beta y_\beta \in V$, and $y_\beta x_\beta \in W$, then $x_\beta^{-1}(x_\beta y_\beta)x_\beta \in x_\beta^{-1}Vx_\beta$ and $x_\beta^{-1}(x_\beta y_\beta)x_\beta = y_\beta x_\beta \in W$. Thus $x_\beta^{-1}Vx_\beta \not\subset U$, and G has inequivalent left and right uniform structures by the above remark.

Conversely, assume that G is a group with inequivalent left and right uniform structures. Then there is a neighborhood U of e such that in every neighborhood V of e exists a $t \in G$ such that $tVt^{-1} \not\subset U$. Then for every neighborhood V of e contained in U, there are $t_V \in G$ and $s_V \in V$ such that $t_V s_V t_V^{-1} \notin U$. Introduce an ordering in the family $\{V\colon V \subset U\}$ of neighborhood of e as follows: for every pair of neighborhood V_1, V_2 of e contained in U, define $V_1 \leqq V_2$ if and only if $V_2 \subset V_1$. Let $x_V = t_V^{-1}$ and $y_V = t_V s_V$, then $\{x_V\}$ and $\{y_V\}$ are nets in G such that the net $\{x_V y_V\}$ converges to e, but e is not a cluster point of the net $\{y_V x_V\}$ since $y_V x_V \notin U$ for each V.

If G is a topological group, and if $N = \overline{\{e\}}$, the closure of $\{e\}$ in G, then G/N is a Hausdorff topological group, called the Hausdorff topological group associated with the topological group G.

It is noted that if H is a normal subgroup of a topological group G with equivalent left and right uniform structures, then the factor group G/H is also a group of such kind. To see this, suppose η is the natural map of G onto G/H and assume U is a neighborhood of H in G/H, then $\eta^{-1}(U)$ is a neighborhood of the identity e in G, and so there is a neighborhood V of e such that $tVt^{-1} \subset \eta^{-1}(U)$ for all $t \in G$. This implies that $t\eta(V)\, t^{-1} \subset U$ for all $t \in G/H$ and G/H has equivalent left and right uniform structures.

COROLLARY. *A topological group G has equivalent left and right uniform structures if and only if its associated Hausdorff topological group G/H has equivalent left and right uniform structures.*

Proof. The necessity is clear as stated above.

For the sufficiency, assume that G/N has equivalent left and right uniform structures but G is not. Then, by the theorem, there are nets $\{x_\alpha\}$ and $\{y_\alpha\}$ in G such that $\{x_\alpha y_\alpha\}$ converges to e, but e is not a cluster point of the net $\{y_\alpha x_\alpha\}$. It follows that $\{x_\alpha N\}$ and $\{y_\alpha N\}$ are nets in G/N such that $\{x_\alpha y_\alpha N\}$ converges to N in G/N. I claim that N is not a cluster point of $\{y_\alpha x_\alpha N\}$. To see this, suppose U were a neighborhood of the identity e in G such that the net $\{y_\alpha x_\alpha\}$ is eventually in the complement of U, and let V be a symmetric neighborhood of e such that $V^2 \subset U$. Then the net $\{y_\alpha x_\alpha N\}$ would eventually be in the complement of VN. This would imply that G/N is a group with inequivalent left and right uniform structures, which contradicts the assumption.

Example. Let G and H be the multiplicative group of positive real numbers. For each number h in H, let T_h be the automorphism of G defined by $T_h(x) = x^h$, for x in G. Since the mapping $(x, h) \to T_h(x)$ of $G \times H$ into G is continuous, the semidirect product $G \otimes H$ using the multiplication $(x, h)(a, b) = (xa^h, hb)$ is a Hausdorff topological group whose identity element is $(1, 1)$. For each positive integer n,

let $x_n = (n, 1)$ and $y_n = (n, n)$. Then $x_n y_n = (n^{1/n}, 1)$ and $y_n x_n = (n, 1)$, and so $\{x_n y_n\}$ converges to $(1,1)$, but $(1,1)$ is not a cluster point of the sequence $\{y_n x_n\}$. Hence, by the above theorem, $G \otimes H$ is a group whose left and right uniform structures are not equivalent. We note that this group is in fact the group of matrices of the form $\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$ for x and y reals.

Reference

1. E. Hewitt and K. A. Ross, Abstract Harmonic Analysis, vol. 1, Academic Press, New York, 1963.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

THE OPPORTUNITIES AND PROBLEMS OF THE TWO-YEAR COLLEGE[1]

G. S. YOUNG, University of Rochester

Almost since the establishment of our country, the American people have believed in education as a basic right. At first this meant the right to a basic literacy, to some years of a primary education. Later, this became the right to a complete elementary education. At the beginning of this century the right was extended to include a secondary education. It is now being changed to a right to universal post-secondary education.

It would be easy to say that the right to a post-secondary education has long been admitted, and to point to the multiplicity of land-grant universities, city colleges, and state colleges (not to mention the private institutions); but these institutions are an answer to a different "right to education." The full American policy on education has included not only the universal right to education up to a certain level but also the right to attempt the next higher level, and the public institutions of higher education have provided that right. Although admission to many of these institutions is quite easy, often requiring no more than a high school diploma, being retained is difficult. In many schools with an open admissions policy, 50, 60, or 70 percent of each freshman class still fail. What had been provided was not post-secondary educa-

[1] Modified from an editorial in *The Two-Year College Mathematics Journal*, 1 (1970), published by Prindle, Weber and Schmidt, Boston.

References

1. R. D. Anderson, Are there too many Ph. D's? this MONTHLY, 77 (1970) 626–641.
2. W. L. Duren, Jr., Are there too many Ph. D's? this MONTHLY, 77 (1970) 641–646.
3. The Mathematical Sciences: A Report, National Academy of Sciences, Publication 1681, Washington, D. C., 1968.
4. G. S. Young, The Ph. D Class of 1951, this MONTHLY, 71 (1964) 787–790.
5. ———, The problems of employment in the mathematical sciences, Notices, Amer. Math. Soc., 18 (1971) 718–722.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before July 31, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.
An asterisk (*) means neither the proposer nor the editors supplied a solution.

2349.* *Proposed by C. S. Ogilvy, Hamilton College*

Find the side of the largest cube that can be wholly contained within the regular tetrahedron of side 1.

E 2350. *Proposed by H. D. Ruderman, Hunter College High School*

A total of n fair coins are flipped and laid in a row. What is the probability that in the row neither the combination HTH nor the combination THT occurs anywhere?

E 2351. *Proposed by Stefan Porubsky, Comenius' University Bratislava, Czechoslovakia*

Let ϕ denote Euler's totient function and let $\tau(n)$ denote the number of divisors of n. Show that

$$\phi(n)\,[\tau(n)]^2 \leqq n^2$$

for all positive integers $n \neq 4$. For what n does equality hold?

E 2352. *Proposed by Marlow Sholander, Case Western Reserve University*

For each positive integer n, define

$$Q_n = \left[1+\frac{1}{n}\right]^{n^2} \frac{n!}{n^n\sqrt{n}}.$$

Show that the sequence $\{Q_n\}$ is monotonely decreasing and find its limit.

E 2353. *Proposed by J. G. Rau, Litton Systems, Culver City, Cal.*

Given two sequences $\{a_1, a_2, \cdots, a_n\}$ and $\{b_1, b_2, \cdots, b_n\}$ of positive real numbers, find the permutation $(j_1, \cdots, j_n)$ of the integers $1, 2, \cdots, n$ for which

$$\sum_{m=1}^{n} \sum_{k=1}^{m} b_{j_m} a_{j_k}$$

is a minimum.

E 2354. *Proposed by L. Carlitz and R. A. Scoville, Duke University*

Let $S = \{1, 2, \cdots, n\}$ and let D_n denote the number of permutations of S with no fixed points (derangements). Let E_n denote the number of even permutations of S with no fixed points. Show that

$$E_n = \binom{n}{2} D_{n-2} - (-1)^n(n-1), \qquad n = 2, 3, \cdots.$$

SOLUTIONS OF ELEMENTARY PROBLEMS

Two Triangle Inequalities

E 1838 [1965, 1129; 1967, 440]. *Proposed by A. Oppenheim, University of Ghana*

Suppose that ABC is an acute-angled triangle; then

$$16\,\Pi \cos^2 A + 4 \sum \cos^2 B \cos^2 C \leqq 1, \tag{1}$$

$$4 \sum \cos^2 B \cos^2 C \leqq \sum \cos^2 A. \tag{2}$$

Equality occurs when ABC is equilateral or right-angled isosceles and in no other case.

II. *Comment and solution by Murray Klamkin, Ford Scientific Laboratory.* By virtue of the weak inequality conditions, ABC can be restricted to non-obtuse triangles rather than acute triangles.

In a personal communication, A. W. Walker has pointed out that there is a flaw in the published solution [1967, 441]. He notes that the solution "derives" and uses the inequality $16\,\Sigma\cos^2 B\cos^2 C \leqq 3$; however, this is invalid—just consider an isosceles right triangle. (By continuity, there exist acute, non-isosceles triangles which violate the inequality.)

We prove (2) of the problem and show how (1) follows from it. By using $2\cos^2 A = 1 + \cos 2A$ and then making the transformations $A' = \pi - 2A$, etc., we see that (2) becomes equivalent (after dropping primes) to the following:

$$(3) \qquad 3\ \Sigma\cos A \geqq 3 + 2\ \Sigma\ \cos B\cos C,$$

where now ABC is an arbitrary triangle. Inequality 6.12 of O. Bottema et al., *Geometric Inequalities*, Nordhoff, Groningen, 1969, states $2R + 5r \geqq h_a + h_b + h_c$. Since $h_a = AH + HD = 2R\cos A + 2R\cos B\cos C$, etc., it follows that

$$2R + 5r \geqq 2R\ \Sigma\ \cos A + 2R\ \Sigma\ \cos B\cos C,$$

and hence $5(1 + r/R) \geqq 3 + 2\ \Sigma\cos A + 2\ \Sigma\cos B\cos C$ which reduces to (3) since $1 + r/R = \Sigma\ \cos A$.

Now, using (2) we establish a stronger inequality than (1), viz.

$$(4) \qquad 16\ \Pi\ \cos^2 A + \Sigma\ \cos^2 A \leqq 1.$$

Since $1 - \Sigma\ \cos^2 A = 2\ \Pi\ \cos A$, (4) is equivalent to

$$(5) \qquad (\Pi\ \cos A)(1 - 8\ \Pi\ \cos A) \geqq 0.$$

But $\Pi\cos A \geqq 0$ since the triangle is non-obtuse and $8\ \Pi\cos A \leqq 1$ by 2.24 of Bottema et al. Thus (5) is established. We note that there is equality in (5) if and only if the triangle is equilateral or a right triangle. This implies that there is equality in (1) if and only if the triangle is equilateral or right isosceles.

Nesting Habits of the Laddered Parenthesis

E 1903 [1966, 666; 1970, 525; 1971, 298]. *Proposed by George Eldredge, El Cerrito, California*

Let an n-ladder of twos, L_n, be defined as follows:

$$L_n = 2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}$$

where there are n twos. Let N_n be the number of distinct integers that can be ob-

tained from L_n by the appropriate insertion of a set of unambiguous nested parentheses. For example, $N_3 = 1$, $N_4 = 2$. Find N_n.

Completion of solution by R. P. Nederpelt, Technical University, Eindhoven, Netherlands. Note that all telescoped n-ladders of twos can be written in the form $L_{i,n} = 2^{2^t}$ where t is an integer not less than $n-1$. Call t the *second exponent* of 2^{2^t}. Then $2^{(L_{i,n})}$ has second exponent 2^t and $(L_{i,n})^2$ has second exponent $t+1$.

We can now apply a lemma due to K. A. Post, *A combinatorial lemma involving a divergence criterion for series of positive terms*, this MONTHLY 77(1970), 1085–1087. It is not hard to see that Post's lemma still holds if $k = 2$ whenever $f(2) > 3$. Take $f(n) = 2^n$; obviously f obeys Post's difference condition (D) and $2^2 > 3$, so that application of the lemma with $k = 2$ completes the solution immediately.

Editor's comment. Completions of the solution were also submitted by G. A. Heuer & C. V. Heuer, R. K. Guy & J. L. Selfridge, and Richard Yates. Guy and Selfridge refer to their paper, *The nesting and roosting habits of the laddered parenthesis*, University of Calgary Research paper no. 127, June 1971, in which they attack also the problem of evaluating n-ladders where the parenthese are not necessarily nested. See also F. Göbel and R. P. Nederpelt, *The number of numerical outcomes of iterated powers*, this MONTHLY, 78 (1971) 1097–1103.

More About Magic Star Polygons

E 2265 [1970, 1106; 1971, 1025]. *Proposed by N. M. Dongre, Sydenham College, India*

Let a regular star polygon be constructed by dividing a circle into n equal parts and by drawing chords joining alternate points of division. Each of the n chords will carry four points of intersection. It is desired to assign the integers $1, 2, \cdots, 2n$ to the $2n$ points of intersection so as to have a magic star polygon (i.e., the sum of the four numbers on each chord is constant; see Problem E 2091 [1968, 557]). Prove that a necessary condition for the existence of a magic star polygon is that $n > 5$. Is this condition sufficient?

II. *Comment by Martin Gardner, Hastings-on-Hudson, New York.*

My *Scientific American* column for December 1965 discussed star polygons and their equivalence to problems involving the magic numbering of various polyhedra skeletons. I cited Henry Ernest Dudeney's enumeration of the number of distinct magic stars for $n = 6, 7$, and 8. (The order 8 star considered by Dudeney is not of the type specified in E 2265. In my column I gave a simple proof that the order 8 star considered by E 2265 has no solution.) Dudeney's results however, were not correct; subsequent computer programs have now established that there are 72 order 7 magic stars and 80 order 6 magic stars. A. Domergue, of Paris, finds

112 order 8 magic stars (of Dudeney's type) and estimates that there are more than 2000 magic stars of order 9. I have briefly summarized these results in my notes on Problems 395 and 396 of Dudeney's *536 Puzzles and Curious Problems*, Scribner's, New York, 1967, pp. 351–352.

HI, HO

E 2282 [1971; 196, 542]. *Proposed by W. J. Blundon, Memorial University of Newfoundland*

For any triangle (other than equilateral) with circumcenter O, incenter I, and orthocenter H, let the angles have measures $\alpha \leqq \beta \leqq \gamma$. Prove

$$(1) \qquad 1 < HO/IO < 3 \quad \text{and} \quad 0 < HI/HO < 2/3$$

$$(2) \qquad 0 < HI/IO < 1 \quad \text{if } \beta > 60^\circ, \; HI = IO \quad \text{if } \beta = 60^\circ,$$

$$1 < HI/IO < 2 \quad \text{if } \beta < 60^\circ,$$

and show that the constant 2 in the last inequality cannot be replaced by a smaller number.

Solution by Anders Bager, Hjørring, Denmark. We shall use the following relations (for non-equilateral triangles) which are all well-known, or at least easily derivable from known relations:

(i) $IO^2 = R^2 - 2Rr$
(ii) $HO^2 = 9R^2 + 8Rr + 2r^2 - 2s^2$
(iii) $HI^2 = 4R^2 + 4Rr + 3r^2 - s^2$
(iv) $R > 2r$
(v) $16Rr - 5r^2 < s^2 < 4R^2 + 4Rr + 3r^2$
(vi) s is $>$, $=$, or $<$ $(R + r)\sqrt{3}$ according as β is $>$, $=$, or $< \pi/3$.

(As usual, R, r and s are the circumradius, inradius and semi-perimeter.)

From (ii) and (v) it follows that

$$R^2 - 4r^2 < HO^2 < 9R^2 - 24Rr + 12r^2;$$

dividing this through by $R^2 - 2Rr > 0$, using (i) and extracting square roots, we have

$$(1 + 2r/R)^{\frac{1}{2}} < HO/IO < (9 - 6r/R)^{\frac{1}{2}}$$

and this implies the first inequality of (1).

The inequality $HI^2/HO^2 < 4/9$, which is equivalent to the second inequality of (1), is seen to be equivalent to $s^2 > 4Rr + 19r^2$ by using (ii) and (iii); this last inequality follows easily from (iv) and the first part of (v).

To prove (2) we see that $HI/HO < 2/3$ implies

$$3HI < 2HO \leqq 2(HO + HI) \text{ and hence } HI/IO < 2.$$

Now, by (i) and (iii) we have that HI^2 is $<$, $=$, or $> IO^2$ according as s^2 is $>$, $=$, or $< 3(R+r)^2$. By (vi) it follows that HI/IO is $<$, $=$, or >1 according as β is $>$, $=$, or $< \pi/3$.

We can show that the constant 2 is best possible by considering the triangle with vertices $(0,\varepsilon)$, $(1,0)$, and $(-1,0)$; by actual computation HI/IO can be made arbitrarily close to 2 by making ε sufficiently small.

In Section 14.14 of *Geometric Inequalities* by Bottema et al. (Groningen, 1969) we find the inequality $HO \geqq HI\sqrt{2}$ with equality if and only if the triangle is equilateral. One is tempted to conclude that $\sqrt{2}$ is best possible, but it follows from the second inequality of the problem that $3/2$ is better. This seeming paradox is easily resolved since, when the triangle is equilateral, $I = H = O$, so that all constants will do.

Also solved by Leon Bankoff, Michael Goldberg, M. G. Greening (Australia), John Leech (Scotland), Simeon Reich (Israel), K. R. S. Sastry (Ethiopia), and the proposer.

A Known Number-Theoretic Result

E2295 [1971, 542]. *Proposed by R. S. Luthar, University of Wisconsin, Janesville*

Suppose that m, n and d $(d>1)$ are arbitrary positive integers. Evaluate $(d^{md}-1, d^{nd}-1)$.

Comment by Bob Prielipp, Wisconsin State University, Oshkosh. In W. Sierpinski, *Elementary Theory of Numbers*, Warsaw, 1964, p. 29, it is shown that if $a>1$, then $(a^m-1, a^n-1) = a^s-1$, where $s=(m,n)$. The present problem is the special case $a = d^d$.

Also solved by 37 other readers.

Editorial Comment. Your editors apologize for including a problem which seems to be well known to everyone else. Nilo Niccolai notes that Nagell, *Introduction to Number Theory*, New York, 1951, has an exercise on p. 42 to find (N^m-1, N^n-1). (No solution is given.) O. H. Fraser remarks that in Faddeev & Sominskii, *Problems in Higher Algebra*, San Francisco, 1965, Exercise 109 is to show that if $(a,b)=1$, then the GCD of x^a-1 and x^b-1 is $x-1$. David Zeitlin notes that it is a result of Lucas (Comptes Rendus, Paris, 82 (1876), 1303–1305) that if $u_n = (a^n-b^n)/(a-b)$, where $(a,b)=1$, then $(u_m, u_n) = u_d$, where $d=(m,n)$. (See L. E. Dickson, *History of the Theory of Numbers*, I, p. 396.)

R. T. Bumby remarks that the result can be extended to algebraic integers of special form, and refers to the following papers: R. D. Carmichael, *On the numerical factors of the arithmetic forms* $\alpha^n \pm \beta^n$, Ann. Math., 15 (1913) 30–70; D. H. Lehmer, *An extended theory of Lucas's functions*, Ann. Math., 31 (1930) 419–448; W. Ljunggren, *On the Diophantine equation* $Ax^4 - By^2 = C$ $(C = 1,4)$, Math. Scand., 21 (1967) 149–158.

Prime Power Polynomials

E2296 [1971, 543]. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

A nonconstant polynomial f with integral coefficients has the property that for

each prime p_i, there exists a prime q_i and an integer m_i such that $f(p_i) = q_i^{m_i}$. Prove that the polynomials contained in $\{x^n\}$, $n = 1, 2, \cdots$, are the only polynomials which possess this property. (This generalizes E 1632 [1964, 795].)

Solution by Allen Stenger, student, Emory University. Suppose first that for some p, it is true that $f(p) = q^m$ where $q \neq p$. By actual computation, we see that q^{m+1} divides $(f(p+sq^{m+1})-f(p))$ for $s=1,2,\cdots$. Since $q|f(p)$ but $q^{m+1}\nmid f(p)$, it follows that $q|f(p+sq^{m+1})$ but $q^{m+1}\nmid f(p+sq^{m+1})$. By Dirichlet's theorem, we can choose s so large that $p+sq^{m+1}$ is prime and also so large that $f(p+sq^{m+1}) > q^m$. By assumption, $f(p+sq^{m+1})$ is a power of a prime; since $q|f(p+sq^{m+1})$ it is clear that $f(p+sq^{m+1}) = q^t$ for some t. Since $q^m < f(p+sq^{m+1}) = q^t$, it follows that $m < t$, so that $q^{m+1}|f(p+sq^{m+1})$. This is a contradiction, and therefore for every prime p, $f(p) = p^m$, where m (possibly) depends on p.

Write $f(x) = a_0 + a_1x + \cdots + a_nx^n$, where $a_n \neq 0$. Evidently, for sufficiently large primes p, $f(p) = p^n$. If we put $g(x) = x^n$, then $g(p) = f(p)$ for sufficiently large primes p, i.e., for infinitely many values. Since f and g are polynomials, it follows that $f = g$.

Also solved by Anders Bager (Denmark), D. Borwein & J. M. Borwein, Robert Breusch, R. T. Bumby, Frederick Carty, R. J. Dickson, Neal Felsinger, Harry Lass, Konrad Victor (Israel), Stanley Wagon, and the proposers.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before July 31, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

5848. *Proposed by A. Smith, Carleton University, Ottawa*

Let m and n be positive coprime integers. Find the number of zeros of the function $z^n + z^m - 1$ which lie inside the unit circle.

5849*. *Proposed by H. D. Ruderman, Hunter College High School, New York City*

For positive integers n, what is the Greatest Lower Bound for $n\,|\sin n|$?

5850. *Proposed by R. K. Tamaki, California State College at Los Angeles*

Let X be metrizable. Prove that X is compact if and only if, for every metric d for X, every open cover $\{U_\alpha\}$ of X has a Lebesgue number $\lambda > 0$ (i.e., we require that each d-ball $B_d(x, \lambda)$ is contained in some U_α).

5851. *Proposed by Douglas Lind, Stanford University*

Is there a bounded sequence of real numbers each translate of which has only

finitely many terms in the Cantor set?

5852. *Proposed by C. H. Kimberling, University of Evansville*

Suppose f carrying $[0,\infty)$ onto $(0,1]$ has alternating derivatives:

$$(-1)^k f^{(k)} \geqq 0, \qquad k = 0,1,\cdots.$$

Prove $g(x) = (1-f(x))/x$ has alternating derivatives on $(0,\infty)$.

5853. *Proposed by Gomer Thomas, University of Washington*

Let x and y be elements of a finite Abelian group G with orders m and n respectively. Let q be the order of $\langle x\rangle \cap \langle y\rangle$, the intersection of the cyclic subgroups generated by x and y. Give the possibilities for the order of xy, in terms of m, n, and q.

SOLUTIONS OF ADVANCED PROBLEMS

The Polynomial $F(x) : F(2\cos\theta) = 2\cos p\theta$

5779 [1971, 203]. *Proposed by G. J. Janusz, University of Illinois*

Let p be an odd prime and $F(X)$ the polynomial with rational coefficients such that $F(2\cos\theta) = 2\cos p\theta$ for all real θ. Let m and n be nonzero integers both relatively prime to p such that $|pm| < n$. Set $f(X) = F(X) - 2pm/n$. Prove

(1) $f(X)$ is irreducible over the rationals.

(2) The roots of $f(X)$ are all real.

(3) The Galois group of $f(X)$ over the rationals is solvable with order dividing $2p(p-1)$.

Solution by John Coolidge, Florida State University. Since

$$\begin{aligned} e^{ip\theta} &= \cos p\theta + i\sin p\theta = (\cos\theta + i\sin\theta)^p, \text{ we have} \\ \cos p\theta &= \sum_{j=0}^{(p-1)/2} \binom{p}{2j} (\cos\theta)^{p-2j}(-1)^j(1-\cos^2\theta)^j \\ &= \sum_{j=0}^{(p-1)/2} a_{2j+1}(\cos\theta)^{2j+1}, \end{aligned}$$

where each a_{2j+1} is an integer, $a_p \equiv 1 \pmod p$, and $a_{2j+1} \equiv 0 \pmod p$ for $0 \leqq j \leqq (p-3)/2$. Therefore

$$2\cos p\theta = \sum_{j=0}^{(p-1)/2} c_{2j+1}(2\cos\theta)^{2j+1},$$

where $c_{2n+1} = a_{2n+1}/2^{2j+1}$, and $F(X) = \sum_{j=0}^{(p-1)/2} c_{2j+1}X^{2j+1}$.

The polynomial $2^p n f(X)$ is in $Z[X]$, and is irreducible over the rational field Q by Eisenstein's Criterion, which is applicable, using the prime p; hence $f(X)$ is

irreducible over Q. For any real number θ, $f(2\cos\theta) = 2(\cos p\theta - m/n)$; hence if θ_0 is the unique real number in the interval $(0, \pi/p)$ such that $\cos p\theta_0 = m/n$, then $\{2\cos(\theta_0 + 2\pi j/p)\}_{j=0}^{p-1}$ is a set of p distinct real roots of $f(X)$. Since $f(X)$ has degree p, it follows that all roots of $f(X)$ are real.

Finally, we prove that the Galois group G of $f(X)$ over the rationals is solvable, with order dividing $2p(p-1)$. Since

$$\begin{aligned}\cos(\theta_0 + 2\pi j/p) &= \cos\theta_0\cos(2\pi j/p - \sin\theta_0\sin(2\pi j/p))\\ &= \tfrac{1}{2}[(\zeta^j + \zeta^{-j})\cos\theta_0 - (\zeta^j - \zeta^{-j})\sin\theta_0],\end{aligned}$$

where $\zeta = e^{2\pi i/p}$, the splitting field K of $f(X)$ over Q is contained in the field $Q(\cos\theta_0, \sin\theta_0, \zeta)$. It is clear that $[Q(\cos\theta_0, \sin\theta_0); Q]$ is p or $2p$, and since $Q(\zeta)/Q$ is normal of degree $p-1$, the degree of ζ over $Q(\cos\theta_0, \sin\theta_0)$ divides $p-1$. Hence $[Q(\cos\theta_0, \sin\theta_0, \zeta)\colon Q]$, and therefore $[K\colon Q]$, divides $2p(p-1)$. To prove that G is solvable, it suffices to prove that $Q(\cos\theta_0, \sin\theta_0, \zeta)$ is solvable by radicals over Q. Since $Q(\zeta)$ is solvable by radicals over Q and since $\sin\theta_0 = (1-\cos^2\theta_0)$, it suffices to prove that $Q(\cos\theta_0)$ is solvable by radicals over Q. Now $m/n = \cos p\theta_0 = (\alpha + \alpha^{-1})/2$, where $\alpha = e^{ip\theta_0}$, so that $Q(\alpha)$ is solvable by radicals over Q since $n\alpha^2 - 2m\alpha + n = 0$. But $\cos\theta_0 = (e^{i\theta} - e^{-i\theta})/2$, where $(e^{i\theta_0})^p = \alpha$; consequently, $Q(\cos\theta_0)$ is solvable by radicals over Q, and this completes the proof.

Also solved by M. G. Greening (Australia), Jack Hart, A. A. Jagers (Netherlands), Takashi Tamura (Japan), and the proposer.

Normal Subgroups of a Torsion Group

5782 [1971, 203]. *Proposed by Jiang Luh, North Carolina State University*

Let G be a torsion group and H be a subgroup of G of index m (finite). Show that if all prime factors of the orders of elements of H are $\geqq m$, then H is a normal subgroup of G.

Solution by Z. Z. Uoiea, University of Utah at Lakeside. Let K be the kernel of the permutation representation T of G on the cosets of H. Then $K \subset H$. If $K \neq H$, then there is an element $x \in H$ whose order with respect to K is a prime p. By assumption, $p \geqq m = \deg(T)$. Since $T(x) \neq 1$, $T(x)$ contains a p-cycle. Since $T(x)$ also fixes a letter (namely H), this is impossible. Hence $K = H$ and H is normal in G.

Also solved by James Alonso, R. S. Castroll (Israel), John Coolidge, A. Drillick, W. E. Everidge III, Neal Felsinger, M. G. Greening (Australia), M. L. Hamilton, C. V. Heuer & G. A. Heuer, A. A. Jagers (Netherlands), Sister Janet Schillinger, University of Northern Colorado Group Theory Class, W. C. Waterhouse, Mark Yu, and the proposer.

Homeomorphic Topologies for the Integers Z

5783 [1971, 304]. *Proposed by D. A. Moran, Michigan State University*

Let Z denote the set of integers, and let p be a fixed prime. For each positive

integer a, define

$$U_a(n) = \{n + \lambda p : \lambda \in Z\}.$$

Then, as is well known $\{U_a(n)\}$ is a basis for some topology $\mathscr{T}_p$ on Z. If $p \neq q$, it is easy to show that $\mathscr{T}_p$ and $\mathscr{T}_q$ are distinct topologies on Z, and that $(Z, +, \mathscr{T}_p)$ and $(Z, +, \mathscr{T}_q)$ are not isomorphic topological groups.

Prove or disprove: $(Z, \mathscr{T}_p)$ and $(Z, \mathscr{T}_q)$ are never homeomorphic topological spaces.

Solution by Don Coppersmith, Massachusetts Institute of Technology. We show that $(Z, \mathscr{T}_2)$ is homeomorphic to $(Z, \mathscr{T}_3)$. Define a mapping from Z onto Z as follows:

(1) If the number is positive or zero, expand it in ternary notation. If negative, a $\sim$ at the left of a ternary number stands for -1 times the appropriate power of 3, e.g., $(\sim 121)_3 = -27 + 9 + 6 + 1 = -11$. The $\sim$ may be used only once in a negative number, and only at the left. (The ambiguity between $\sim$ and ~ 2, e.g., $(\sim 2121)_3 = -11$, will be unimportant. For definiteness, suppress all 2's immediately to the right of $\sim$.)

(2) Make the substitutions: $\sim \to \sim$, $0 \to 0$, $1 \to 01$, $2 \to 11$.

(3) Translate the resulting binary number back to Z, interpreting the $\sim$ as -1 times the appropriate power of 2. (Notice that the equivalent ternary forms $\sim$ and ~ 2 become the equivalent binary forms $\sim$ and ~ 11, so, as indicated, the ambiguity is unimportant.)

Now $f: Z \to Z$ is a homeomorphism. First, f is one-one and onto, since an inverse function exists (the process above is reversible, except in step 2 care must be taken to block the binary number from the right and possibly supply a 0 on the left of a positive number, or a 1 to the right of a $\sim$ in a negative number). Second, f is continuous. Notice that basis sets are defined by the right-most collection of bits, which is consistent with our usage of $\sim$. I.e., $U_a(n)$ is the set of numbers whose p-ary expansions have the rightmost a bits identical with those of n, assuming the $\sim$ (if any) is replaced by $\sim (p-1)(p-1)(p-1)$ until the $\sim$ is forced out of the rightmost a bits, and if the number is positive, leading 0's are supplied to fill out the a bits. Then the inverse image of a basis set of $\mathscr{T}_2$ is either a basis set of $\mathscr{T}_3$ or the union of two basis sets of $\mathscr{T}_3$, and is in either case open.

Similarly, f^{-1} is continuous, since the image of a basis set of $\mathscr{T}_3$ is a basis set of $\mathscr{T}_2$.

The proof generalizes.

Also solved by D. P. Robbins, W. C. Waterhouse, and Mark Yu.

Editorial Note. Waterhouse gives a simple solution to the problem using the fact that every denumerable metric space without isolated points is homeomorphic to the rationals. See A. Wilansky, *Topology for Analysis*, page 112.

Quadrature for functions in $C^{(5)}$

5784 [1971, 304]. *Proposed by Anon, Erewhon-upon-Wabash*

Let $x_0 = a$, $x_1 = a + h$, $x_2 = a + 2h$, $x_3 = a + 3h$. Prove the existence of unique polynomials $u(x)$, $v(x)$, $w(x)$ of degree 5 such that

$$\int_{x_0}^{x_1} u(x)f^{(5)}(x)\,dx + \int_{x_1}^{x_2} v(x)f^{(5)}(x)\,dx + \int_{x_2}^{x_3} w(x)f^{(5)}(x)\,dx$$

$$= 44\int_{x_0}^{x_1} f(x)\,dx + 152\int_{x_1}^{x_2} f(x)\,dx + 44\int_{x_2}^{x_3} f(x)\,dx$$

for each $C^{(5)}$ function f which vanishes at x_0, x_1, x_2, x_3.

Solution by G. L. Isaacs, Lehman College of the City University of New York. The substitution $x = x_0 + h(y + 3)/2$ in the integrals shows that unique polynomials exist to satisfy the given relation if and only if unique polynomials exist for the particular case $x_0 = -3$, $h = 2$. In this case, we put $x = -t$ in the integrals; then, since $g(t) = f(-t)$ satisfies the same conditions as $f(t)$, we see that, if they exist, the polynomials u, v, w must satisfy $u(-x) = -w(x)$, $v(-x) = -v(x)$, so that they must be of the form

$$v = k(ax^5 + bx^3 + cx),$$

$$u = k(x^5 + dx^4 + ex^3 + fx^2 + gx + h),$$

$$w = k(x^5 - dx^4 + ex^3 - fx^2 + gx - h).$$

Successive integration by parts of the integrals involving $f^{(5)}$ yields a sum in terms of $f^{(4)}(3), f^{(3)}(3), f^{(2)}(3), f^{(1)}(3), f^{(4)}(1), f^{(3)}(1), f^{(2)}(1), f^{(1)}(1)$, similar expressions with the points 3, 1 replaced by -3, -1, and

$$\text{(A)} \qquad k(-120)\left[\int_{-3}^{-1} f + a\int_{-1}^{1} f + \int_{1}^{3} f\right].$$

Putting the coefficients of the eight terms mentioned equal to 0 gives eight linear equations for the eight unknowns a through h (the other eight equations are identical), viz.

$$81d + 27e + 9f + 3g + h + 243 = 0, \qquad 108d + 27e + 6f + g + 405 = 0,$$

$$108d + 18e + 2f + 540 = 0, \qquad 72d + 6e + 540 = 0,$$

$$a + b + c - d - e - f - g - h - 1 = 0, \quad -5a - 3b - c + 4d + 3e + 2f + g + 5 = 0,$$

$$20a + 6b - 12d - 6e - 2f - 20 = 0, \qquad -60a - 6b + 24d + 6e + 60 = 0.$$

These have a unique solution, and in particular $a = 38/11$. Finally (A) becomes

(B) $$44\int_{-3}^{-1} f + 152 \int_{-1}^{1} f + 44\int_{1}^{3} f$$

if $k = -44/120$; and by taking $f = 0$ in $[1,3]$ and in $[-3,-1]$, and positive in $(-1,1)$, we see that (A) agrees with (B) only if $k = -44/120$. Thus the polynomials are unique.

Also solved by Harley Flanders, G. A. Heuer, Saint Olaf College Students, and E. T. Wong.

Note. The original statement of the problem had 155 in place of the correct 152. This was noted by all the solvers.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Calculus.* By. H. Flanders, R. Korfhage, and J. Price. Academic Press, New York, 1970. 986 pp. $ 13.95. (Telegraphic Review, March 1970.)

Calculus teaches the calculus by examples, realistic advice, and plenty of practical experience. "Our presentation is informal ... we omit technicalities that almost never occur in practice ... we are always result-oriented and insist on explicit numerical answers ... occasionally we allow ourselves the liberties of circular arguments."

Calculus is readable; many of our students read it and got a decent working knowledge of differentiation, integration, functions, and numbers. This is the fundamental mathematical experience which the great theories of the calculus aim to describe. Unfortunately, the current practice in teaching is to exhibit only specimens of this experience as an excuse for introducing some formal theory. "Since $x^2 = 2$ has no rational solution, we shall now introduce the following five axioms..." *Calculus* rejects this practice. Its great virtue is to present the mathematics relatively "theory-free", i.e., against the background of the student's actual mathematical world. Later one may make theories. Although *Calculus* appears to be aimed primar-

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 5

CONTENTS

(Continued on inside cover)

MAY 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the Monthly. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 11 months, Math. Notes 10 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 7 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to Harley Flanders, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

Published by the Association at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

GEOMETRIC PROBLEMS IN COMPLEX ANALYSIS

T. H. MacGREGOR, State University of New York at Albany

1. Introduction. This paper discusses certain geometric ideas and problems which occur in complex analysis. It is not surprising that geometry bears on this field since a function $w = f(z)$ may be interpreted as a mapping of one set in the z-plane to another in the w-plane. The very problem asked about a function may be a geometric one. For example, one can ask the general question: if a function is analytic on a given set and has other prescribed properties what can be said about the geometry of the range?

An example of this is the classical Koebe $\frac{1}{4}$-theorem [**15**, p. 3]: if f is analytic and one-to-one for $|z| < 1$, $f(0) = 0$ and $f'(0) = 1$ then the range of f contains the disk $|w| < \frac{1}{4}$. Another example is the fact that the range of such a function has an area of at least π. We shall prove this last assertion as well as several other geometric statements about the range of various functions analytic in the open unit disk.

Perhaps a deeper way in which geometry affects complex analysis is the intuitive insight it affords. This can even be crucial in solving an analytic problem. A simple illustration is the classical area theorem for univalent functions [**15**, p. 2]. The evident geometric fact that a certain area is non-negative yields an analytic statement by expressing that area in a series depending on the given function.

The significance of geometric ideas and problems in complex analysis is what is suggested by the term "geometric function theory." Of course, geometric ideas also occur in real analysis, as in calculus through the interpretation of a function $y = f(x)$ by its graph in the $x - y$ plane. But geometry has had a much greater impact in complex analysis and it is a very fundamental aspect of its vitality. Through this paper we hope to give the reader a sense of the kind of geometric arguments made in this area of complex analysis. We also try to indicate how geometric problems may be attacked as well as what analytic tools may be useful.

The results we discuss are quite striking and easy to interpret. A number of geometric ideas and constructions are presented and various appeals are made to the geometric intuition of the reader. Some background in complex analysis is needed but this has been kept to a minimum. The few more advanced results used are presented as clearly as possible and their relationship to the development is easy to understand.

We begin with a presentation of the Open Mapping Theorem for analytic functions. This is not proved but is taken as a convenient starting point for our development. The Maximum Modulus Principle is an immediate (geometric) consequence of this, and then Schwarz's Lemma and the Principle of Subordination are obtained.

Thomas MacGregor received his University of Penn. Ph. D. in 1961 under O. H. Alisbah. He has held positions at Rutgers Univ., Camden, Lafayette College, and presently S. U. N. Y. at Albany. He has published extensively in complex function theory. *Editor.*

Subordination plays a special role in this paper. It is an extraordinarily useful idea for solving geometric problems about analytic functions, and we give numerous applications of this principle.

The ideas and results discussed here are not new and can be found in various forms in books and research articles. Appropriate references are for the most part saved until the last section of the paper.

2. Open Mapping Theorem, Maximum Modulus Theorem, Schwarz's Lemma. Recall that a **neighborhood** of a complex number z_0 is any open disk with center at z_0, that is, a set of the form $\{z: |z - z_0| < r\}$ for some $r > 0$. A set $\mathcal{O}$ of complex numbers is called **open** if each point of $\mathcal{O}$ has a neighborhood contained in $\mathcal{O}$.

We begin with the Open Mapping Theorem, which is the assertion that each non-constant analytic function is an open mapping. More precisely, if f is analytic and nonconstant on a domain (an open, connected set) D and if $\mathcal{O}$ is any open subset of D, then $f(\mathcal{O})$ is open. Briefly, f maps open sets onto open sets. This theorem is not proved here and it is ordinarily obtained from the argument principle.

The Open Mapping Theorem affords a simple geometric proof of the Maximum Modulus Theorem. This is the fact that if a function f is analytic and non-constant on a domain D, then f cannot assume a maximum modulus in D. In other words, there is no point z_0 in D such that $|f(z)| \leqq |f(z_0)|$ for all z in D. This immediately follows from the Open Mapping Theorem since, in particular, $f(D)$ is open and, thus, to each point $w_0 = f(z_0)$ in $f(D)$ there is a neighborhood N of w_0 contained in $f(D)$. It is clear that some points $w = f(z)$ in N satisfy $|w| > |w_0|$. Such points are illustrated by the shaded region in Figure 1 where the circle $|w| = |f(z_0)|$ and the boundary of N are pictured.

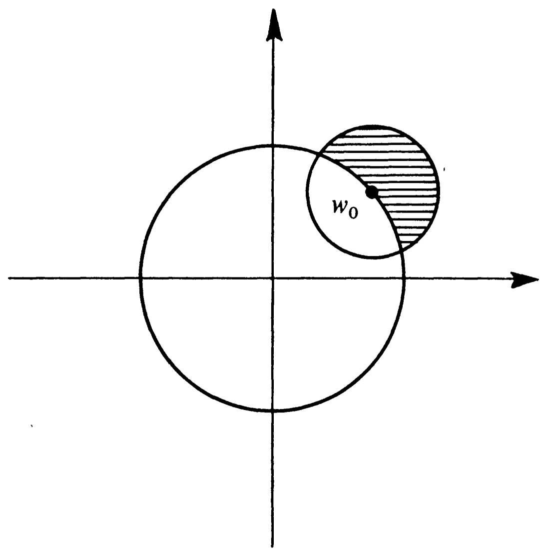


FIG. 1

Similar properties of (non-constant) analytic functions are easy consequences of the fact that they are open mappings. For example, the function $\operatorname{Re} f(z)$ cannot assume a maximum (or minimum) on a domain D if f is analytic and non-constant there. A picture indicating this fact is illustrated by Figure 2. The disk in the figure is a neighborhood of the point $w_0 = f(z_0)$ lying in $f(D)$. All points z corresponding to the points $w = f(z)$ which are in the shaded region satisfy $\operatorname{Re} f(z) > \operatorname{Re} f(z_0)$.

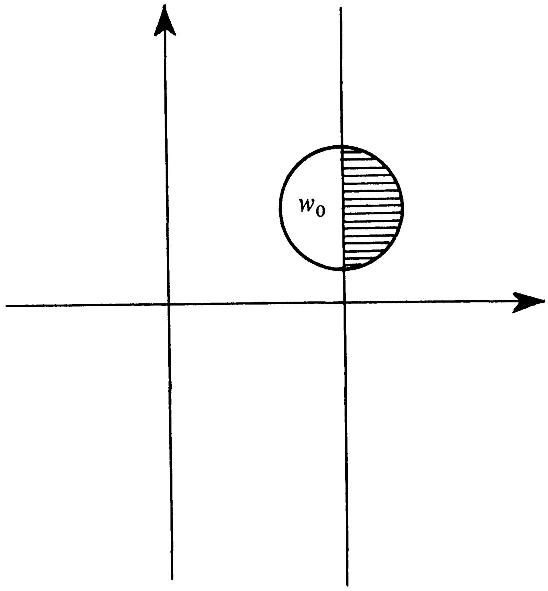


FIG. 2

A more positive statement concerning the Maximum Modulus Theorem occurs when f is analytic in a bounded domain D and continuous on the closure of D. In this case f must achieve a maximum modulus on the closure of D, as the function $|f|$ is continuous on a compact set. But if f is non-constant on D this maximum cannot occur at a point in D, and thereby occurs on the boundary of D. A simple situation in which this is useful is when f is analytic for $|z| < 1$, and it then implies that

$$(1) \qquad \max_{|z| \leq r} |f(z)| = \max_{|z| = r} |f(z)|$$

for each r, where $0 < r < 1$.

A simple application of the Maximum Modulus Theorem concerns the geometric problem: What is the maximum of the products of the four distances between a variable point in a square and each vertex of the square? It is interesting to note the apparent temptation of falsely assuming that the maximum occurs at the center of the square. If we let a, b, c, and d denote the complex numbers giving the vertices of the square, then the problem is the same as maximizing $|f|$, where $f(z) = (z - a)(z - b)(z - c)(z - d)$, and z varies over the square. This must occur at

some point on the perimeter of the square because of the Maximum Modulus Theorem. After this simplification the problem can be solved using some elementary calculus.

Now we discuss Schwarz's lemma, which is the following statement. If the function ϕ is analytic for $|z|<1$ and satisfies $|\phi(z)|<1$ and $\phi(0)=0$, then $|\phi(z)| \leqq |z|$ for each $z(|z|<1)$ and $|\phi'(0)| \leqq 1$. To prove this let the power series representation for ϕ be given by $\phi(z)=\sum_{n=0}^{\infty} a_n z^n$ for $|z|<1$. Then $a_0=\phi(0)=0$, so that we may write $\phi(z)=z\omega(z)$, where $\omega(z)=\sum_{n=0}^{\infty} a_{n+1}z^n$ is analytic for $|z|<1$. Let z_0 satisfy $0<|z_0|<1$ and choose r so that $|z_0| \leqq r<1$. Then,

$$(2) \qquad |\omega(z_0)| \leqq \max_{|z| \leqq r} |\omega(z)| = \max_{|z|=r} |\omega(z)| = \max_{|z|=r} \left|\frac{\phi(z)}{z}\right| \leqq \frac{1}{r}.$$

Since $|\omega(z_0)| \leqq 1/r$ for each r, where $|z_0| \leqq r<1$, this implies that $|\omega(z_0)| \leqq 1$. Thus, ω satisfies $|\omega(z)| \leqq 1$ for $|z|<1$, and this is the same as $|\phi(z)| \leqq |z|$. Since $\phi'(0)=a_1=\omega(0)$ and $|\omega(0)| \leqq 1$, this completes the proof of Schwarz's lemma.

We note additionally that if $|\phi(z_0)|=|z_0|$ for some z_0 with $0<|z_0|<1$ then ω would achieve its maximum modulus at z_0 and so must be constant; that is, ϕ takes on the form $\phi(z)=\varepsilon z$, where $|\varepsilon|=1$. The same argument (at $z_0=0$) shows that these functions $\phi(z)=\varepsilon z$ are the only ones for which $|\phi'(0)|=1$.

The conclusion $|\phi(z)| \leqq |z|$ in Schwarz's lemma expresses a specific restriction on $\phi(z)$ in terms of z; that is, the mapping $z \to \phi(z)$ does not increase modulus. In particular, it implies that if z varies in any subset of the disk $|z| \leqq r$ $(0<r<1)$ then the values of $\phi(z)$ also lie in that disk. Similarly, the conclusion $|\phi'(0)| \leqq 1$ may be viewed as expressing the fact that of all such functions ϕ the functions $\phi(z)=\varepsilon z$, where $|\varepsilon|=1$, achieve the largest (modulus) derivative at $z=0$.

An important application of Schwarz's lemma concerns the uniqueness question for conformal mappings. For example, suppose that ϕ is a one-to-one analytic mapping of $|z|<1$ onto $|w|<1$ so that $z=0$ corresponds to $w=0$. Then by Schwarz's lemma $|\phi'(0)| \leqq 1$. Since ϕ is one-to-one, ϕ has an analytic inverse η that also satisfies Schwarz's lemma. Therefore, $|\eta'(0)| \leqq 1$, which is the same as $|\phi'(0)| \geqq 1$ since $\eta'(0)=1/\phi'(0)$. Thus, we must have $|\phi'(0)|=1$, which is only possible for the functions $\phi(z)=\varepsilon z$, where $|\varepsilon|=1$. Our conclusion is that the only one-to-one analytic maps of $|z|<1$ onto $|w|<1$ so that 0 corresponds to 0 are these functions. If we also demand that ϕ has a real positive derivative at $z=0$ then ϕ is uniquely determined to be $\phi(z)=z$. It is this kind of consideration which yields the uniqueness statement for the Riemann Mapping Theorem [**32**, p. 175].

3. The principle of subordination. A function f is called **univalent** (or **schlicht**) in D if it is one-to-one there, that is, if f takes on no value more than once in D. Expressed differently, if $f(z_1)=f(z_2)$ with z_1 and z_2 in D, then $z_1=z_2$.

Let f and g be two functions analytic for $|z|<1$ with ranges F and G, respectively, and suppose that $F \subset G$. Further let g be univalent for $|z|<1$ and let $f(0)=g(0)$.

These several assumptions are expressed by saying that f is subordinate to g for $|z| < 1$.

Under these conditions g has an analytic inverse g^{-1} and (as $F \subset G$) the function $\phi = g^{-1}(f)$ is analytic for $|z| < 1$. Also $|\phi(z)| < 1$ for $|z| < 1$ and $\phi(0) = 0$ as $f(0) = g(0)$. Thus ϕ satisfies Schwarz's lemma. We may write $f(z) = g(\phi(z))$ and, in particular, find that $f'(0) = g'(0)\phi'(0)$. The inequality $|\phi'(0)| \leqq 1$ implies that

$$|f'(0)| \leqq |g'(0)|. \tag{3}$$

From a knowledge of which functions ϕ satisfy $|\phi'(0)| = 1$ we see that equality in (3) can occur only if $f(z) = g(\varepsilon z)$, where $|\varepsilon| = 1$. Thus, inequality (3) expresses the fact that the maximum of the modulus of the derivative at 0 of all functions subordinate to g occurs exactly for those functions obtained from g by rotations in z.

Subordination is equivalent to the relation $f(z) = g(\phi(z))$, where ϕ satisfies Schwarz's lemma. Therefore, ϕ satisfies $|\phi(z)| \leqq |z|$ for each z ($|z| < 1$). In particular, if $|z| \leqq r$, where $0 < r < 1$, then $|\phi(z)| \leqq r$. Because $f(z) = g(\phi(z))$ this implies that the image of $|z| \leqq r$ under f is a subset of the image of $|z| \leqq r$ under g. This result is usually called Lindelöf's principle. It may be expressed by saying that if f is subordinate to g for $|z| < 1$ then f is subordinate to g for $|z| < r$ for each r, $0 < r < 1$.

The use of the term the **principle of subordination** refers to either inequality (3), the relation of $f(z) = g(\phi(z))$, or Lindelöf's principle. Schwarz's lemma may be thought of as a special case of this principle, where G is the open unit disk and $g(z) = z$.

4. Applications of the principle of subordination. Our first application of subordination concerns starlike mappings. A set D is called **starlike** with respect to the point w_0 if to each point w in D the line segment with endpoints w and w_0 is also in D. A function g analytic for $|z| < r$ is called **starlike** for $|z| < r$ if g is univalent for $|z| < r$, $g(0) = 0$, and the range of g is starlike with respect to the origin. Let D denote the range of a function g starlike for $|z| < 1$. To each point w in D all the points tw, $0 \leqq t \leqq 1$, also belong to D. This is the same as saying tg is subordinate to g for $|z| < 1$ for each t, $0 \leqq t \leqq 1$. Consequently tg becomes subordinate to g for $|z| < r$ $(0 < r < 1)$ for each t, $0 \leqq t \leqq 1$. In other words the image of $|z| < r$ under g is starlike. Briefly, if g is starlike in $|z| < 1$ then g is starlike in $|z| < r$ $(0 < r < 1)$. This is the initial step in an argument that shows that starlike mappings are characterized by the criteria $\operatorname{Re}\{zg'(z)/g(z)\} > 0$ for $|z| < 1$ [**32**, p. 221].

A second application of the **principle of subordination** concerns functions which are analytic for $|z| < 1$ and satisfy $\operatorname{Re} f(z) > 0$. Also assume that $f(0) = 1$ and let $\mathscr{P}$ denote the family of all such functions. Notice that the analytic function $g(z) = (1 + z)/(1 - z)$ satisfies $g(0) = 1$ and maps $|z| < 1$ one-to-one onto the domain $\operatorname{Re} w > 0$. Part of this assertion follows from the computation

$$\text{(4)} \qquad \operatorname{Re}\frac{1+z}{1-z} = \frac{1-|z|^2}{|1-z|^2}.$$

It is more interesting to note that the relation $w = (1+z)/(1-z)$ may be uniquely solved for z to get $z = (w-1)/(w+1)$ and then it is geometrically clear that points w corresponding to $\operatorname{Re} w > 0$ associate with points z with $|z| < 1$ as w is closer to 1 than to -1. The various properties of this function g shows that the family $\mathscr{P}$ consists of exactly those functions that are subordinate to g for $|z| < 1$.

Since $g'(0) = 2$ inequality (3) shows that $|f'(0)| \leqq 2$ for every function f in $\mathscr{P}$. If we express f in a power series

$$\text{(5)} \qquad f(z) = \sum_{n=0}^{\infty} a_n z^n,$$

then $a_0 = f(0) = 1$ and $|f'(0)| \leqq 2$ is the same as $|a_1| \leqq 2$. Without much more effort it is possible to use the principle of subordination to deduce that $|a_n| \leqq 2$ for $n = 1, 2, 3, \cdots$ [**11**, p. 199]. Here we merely note that g has the power series expansion $g(z) = 1 + \sum_{n=1}^{\infty} 2z^n$.

If $0 < r < 1$ then the image of the disk $|z| \leqq r$ under $g(z) = (1+z)/(1-z)$ is the disk

$$\text{(6)} \qquad \left| w - \frac{1+r^2}{1-r^2} \right| \leqq \frac{2r}{1-r^2}.$$

This is a simple mapping problem which can also be solved as follows. Since g is a linear transformation (and $0 < r < 1$) the disk $|z| \leqq r$ must be mapped onto some disk by g. What disk it is can be determined by the facts that

$$g(r) = \frac{1+r}{1-r} \text{ and } g(-r) = \frac{1-r}{1+r}$$

are real and so, by conformality, the boundary of that disk must intersect $g(r)$ and $g(-r)$ perpendicular to the real axis; that is, the disk is that one having as a diameter the segment on the real axis with the end points $(1-r)/(1+r)$ and $(1+r)/(1-r)$. This is the same disk mentioned in equation (6). According to Lindelöf's Principle, the image of $|z| \leqq r$ under any function f in $\mathscr{P}$ is a subset of that disk. Thus, if $f \in \mathscr{P}$ and $|z| \leqq r$, then

$$\text{(7)} \qquad \left| f(z) - \frac{1+r^2}{1-r^2} \right| \leqq \frac{2r}{1-r^2}$$

[**32**, p. 173, problem 11]. This geometric restriction on the numbers $f(z)$ yields various more special results for functions in $\mathscr{P}$ such as the following:

$$\text{(8)} \qquad |f(z)| \leqq \frac{1+|z|}{1-|z|},$$

$$\operatorname{Re} f(z) \geqq \frac{1-|z|}{1+|z|}, \tag{9}$$

$$|\operatorname{Im} f(z)| \leqq \frac{2|z|}{1-|z|^2}, \tag{10}$$

$$|\arg f(z)| \leqq \sin^{-1}\left(\frac{2|z|}{1+|z|^2}\right). \tag{11}$$

For example, (11) follows by determining the maximum angle of all numbers w lying in that disk.

Results like (8) through (11) are called "distortion theorems" since they give restrictions on the growth of quantities depending on f. The family $\mathscr{P}$ gives a good example of a case where the properties of a function can be simply interpreted through subordination so as to yield such analytic results.

A word is in order about the hypothesis (or "normalization") $f(0)=1$. If f is analytic for $|z|<1$ and satisfies $\operatorname{Re} f(z)>0$ and if $f(0)=a+ib$, then $F(z)=(f(z)-ib)/a$ belongs to $\mathscr{P}$. Results known for $\mathscr{P}$ therefore yield information on f. For example, (8) applied to F shows that

$$|f(z)-i\operatorname{Im} f(0)| \leqq \frac{1+|z|}{1-|z|}\operatorname{Re} f(0), \tag{12}$$

or the earlier inequality $|a_n| \leqq 2$ applied to F yields $|b_n| \leqq 2\operatorname{Re} b_0$, if f has the power series $f(z)=\sum_{n=0}^{\infty} b_n z^n$. In general, results obtained for $\mathscr{P}$ can be easily expressed in terms of f with $f(0)$ arbitrary. Making the normalization $f(0)=1$ has the added advantage of expressing results more simply.

Our third application of subordination is the following theorem: *If f is analytic for $|z|<1$, satisfies $f(0)=0$ and $f'(0)=1$, and if the range D of f is convex then D contains the disk* $|w|<\frac{1}{2}$. This is called the $\frac{1}{2}$-covering theorem and is ordinarily associated with univalent, convex mappings. Recall that a convex set D is defined as having the property that if w_1 and w_2 belong to D then the line segment w_1w_2 also belongs to D.

If the above function f does not contain a point c in its range D, we need only show that $|c| \geqq \frac{1}{2}$. To do this we may assume that c has the least modulus of all such numbers (such a number c exists as the complement of D is closed). This point c belongs to the boundary of D, and, as D is convex, there is a support line to D through c, that is, a line L through c exists so that D is a subset of one of the open half-planes determined by L (and in this case it is the one containing $w=0$). We claim that L must be perpendicular to the line through c and 0. Otherwise there are points on that line which are simultaneously in the complement of D and have a smaller modulus than c. This impossibility is illustrated in Figure 3 by the points on the open line segment cd. The circle $|w|=|c|$ is drawn and L^* represents the actual position

L must have. Therefore, the range of f is a subset of some open half-plane which contains the origin and whose bounding line has the distance $|c|$ from the origin.

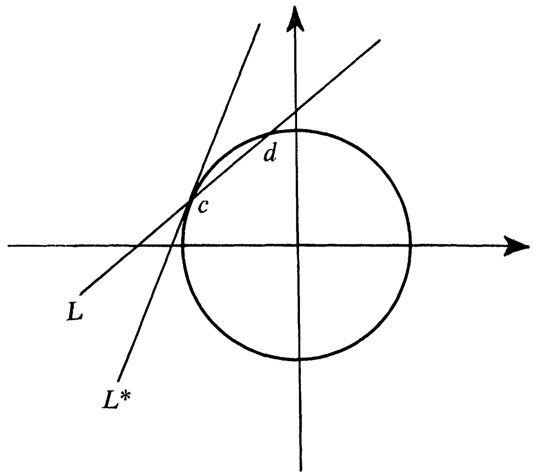


FIG. 3

These properties of f are precisely a subordination relation. Notice that since

$$\frac{z}{1-z} = \frac{1}{2}\left(\frac{1+z}{1-z}\right) - \frac{1}{2} \tag{13}$$

we can deduce that $w = z/(1-z)$ maps $|z| < 1$ one-to-one onto the half-plane $\operatorname{Re} w > -\frac{1}{2}$ due to the corresponding properties of the mapping $w = (1+z)/(1-z)$. Thus, f is subordinate to the function

$$g(z) = 2\varepsilon|c|\frac{z}{1-z}$$

by a suitable choice of the complex number ε with $|\varepsilon| = 1$, as the number ε merely serves to rotate a half-plane into the appropriate direction. Because of (3), $|f'(0)| \leqq |g'(0)|$, which is the same as $|c| \geqq \frac{1}{2}$ due to the normalization $f'(0) = 1$. This proves the $\frac{1}{2}$-covering theorem. We note further that since $w = z/(1-z)$ maps $|z| < 1$ onto $\operatorname{Re} w > -\frac{1}{2}$ there is no number $r > \frac{1}{2}$ such that all functions f have ranges D covering $|w| < r$. We express this by saying that the $\frac{1}{2}$-covering theorem is "sharp."

The $\frac{1}{2}$-covering theorem may be thought of as an improvement of the Koebe $\frac{1}{4}$-theorem quoted in the introduction in the sense that the number $\frac{1}{4}$ is increased to $\frac{1}{2}$ through the added hypothesis that the range is convex. Here again are illustrations of results which may be expressed without the normalizations ($f(0) = 0$ and $f'(0) = 1$), but without them some simplicity is lost. For example, if f is merely

analytic and univalent for $|z| < 1$ then

$$F(z) = \frac{f(z) - f(0)}{f'(0)} \tag{14}$$

also has these properties and additionally satisfies $F(0) = 0$ and $F'(0) = 1$. Applying the Koebe $\frac{1}{4}$-theorem to F shows that the range of f contains the open disk with center $f(0)$ and with radius $\frac{1}{4}|f'(0)|$.

5. Some geometric inequalities for analytic functions. We shall determine certain geometric properties of the range of a function analytic for $|z| < 1$. They concern such quantities as the area of the range of f and the length of the boundary of that range.

Let f be analytic for $|z| < 1$ and let $D(r,f)$ denote the image of $|z| < r$ under f, where $0 < r < 1$. Also let $A(r,f)$ denote the area covered by the map $w = f(z)$ for $|z| < r$, and let $L(r,f)$ denote the length of the curve traced out by the map $w = f(z)$ for $|z| = r$. For example, if $f(z) = z^n$ and n is a positive integer, then $A(r,f) = n\pi r^{2n}$ as f maps the disk $|z| < r$ onto the disk $|w| < r^n$ covered exactly n times. In this example, as z describes the circle $|z| = r$ once then $f(z)$ winds around the circle $|w| = r^n$ precisely n times and consequently $L(r,f) = n2\pi r^n$. A simple situation occurs when f is univalent for $|z| < r$ as then $A(r,f)$ is the area of the set $D(r,f)$ and $L(r,f)$ is the length of the boundary of $D(r,f)$. We consider the problem of finding the least values for $A(r,f)$ and $L(r,f)$ given that f has the normalization $f'(0) = 1$.

These problems are solved by first finding analytic expressions for $A(r,f)$ and $L(r,f)$. The quantity $|f'(z_0)|$ represents the local magnification of length at $z = z_0$ given by the mapping $z \to f(z)$. Likewise, $|f'(z_0)|^2$ represents the local two dimensional magnification of this mapping. It is, therefore, expected that

$$A(r,f) = \iint\limits_{|z|<r} |f'(z)|^2 dx\, dy \qquad (z = x + iy). \tag{15}$$

One more accurately proves (15) by considering the Jacobian,

$$J = J(u,v) = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}, \tag{16}$$

of the transformation $u = u(x,v)$, $v = v(x,y)$, where $f(z) = u + iv$ and $z = x + iy$. Note that

$$J = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 \tag{17}$$

due to the Cauchy-Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}, \tag{18}$$

and since $f'(z) = (\partial u/\partial x) + i(\partial v/\partial x)$ this shows that $J = |f'(z)|^2$. Relation (15) then follows from the result in advanced calculus that expresses the area of the range of such a (smooth) transformation by integration of the Jacobian over the domain.

Let the power series representation for f be given by

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \text{ for } |z| < 1. \tag{19}$$

The integral in (15) may be expressed in polar coordinates with $z = \rho e^{i\theta}$. If this is written as an iterated integral, then since

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1} \text{ by writing } |f'(z)|^2 = f'(z)\overline{f'(z)},$$

we find that (15) becomes

$$A(r,f) = \int_0^r \left\{ \int_0^{2\pi} \left[\sum_{n=1}^{\infty} n a_n \rho^{n-1} e^{i(n-1)\theta} \right] \left[\sum_{m=1}^{\infty} m \bar{a}_m \rho^{m-1} e^{-i(m-1)\theta} \right] \rho d\theta \right\} d\rho. \tag{20}$$

Consider multiplying the two series in the brackets to form a new series in powers of $e^{i\theta}$ and then integrate that series term by term over the interval $0 \leqq \theta \leqq 2\pi$. Because

$$\int_0^{2\pi} e^{ik\theta} d\theta = 0 \tag{21}$$

for every non-zero integer k, the only contributions after this integration will come from terms associated with the case $m = n$. This produces the formula

$$A(r,f) = \int_0^r \left\{ 2\pi \sum_{n=1}^{\infty} n^2 |a_n|^2 \rho^{2n-1} \right\} d\rho. \tag{22}$$

Another term-by-term integration shows that

$$A(r,f) = \pi \sum_{n=1}^{\infty} n |a_n|^2 r^{2n}. \tag{23}$$

This derivation of (23) depended on two term-by-term integrations which can be justified by appropriate appeals to uniform convergence of the series. The earlier multiplication of the two series in θ to get, say, the Cauchy product is justified by absolute convergence of the series.

If f satisfies $f'(0) = 1$, then (23) implies that

$$A(r,f) \geqq \pi |a_1|^2 r^2 = \pi |f'(0)|^2 r^2 = \pi r^2. \tag{24}$$

Thus we have obtained the geometric inequality

$$A(r,f) \geqq \pi r^2. \tag{25}$$

We also see that $A(r,f)$ can equal πr^2 only if $a_n = 0$ for $n = 2, 3, \cdots$, that is, when f has the form $f(z) = a_0 + z$. Since πr^2 is the area of the disk $|z| < r$ inequality (25) asserts that with regard to disks $|z| < r$, f is a mapping that does not decrease area.

Formula (23) and inequality (25) are also valid when $r = 1$. That is, if A denotes the area covered by the mapping $z \to f(z)$ for $|z| < 1$ then

$$A = \pi \sum_{n=1}^{\infty} \cdot n |a_n|^2, \tag{26}$$

where the numbers $\{a_n\}$ are given by the power series development of f. Formula (26) is also interpreted to mean A is finite if and only if the series converges. To prove these assertions let $\{r_k\}$ denote an increasing sequence of real numbers so that $0 < r_k < 1$ and $r_k \to 1$ as $k \to \infty$. Then $D(r_k,f)$ is an increasing sequence of sets, that is, $D(r_k,f) \subset D(r_{k+1},f)$, and $D = \bigcup_{k=1}^{\infty} D(r_k,f)$. Therefore, (because of a basic theorem about two-dimensional Lebesgue measure) $A(r_k,f) \to A$ as $k \to \infty$. If, for example the series $\sum_{n=1}^{\infty} n|a_n|^2$ converges then the function

$$B(r) = \pi \sum_{n=1}^{\infty} n |a_n|^2 r^{2n}$$

is continuous for $0 \leqq r \leqq 1$, and, therefore,

$$A(r_k,f) = B(r_k) \to B(1) = \pi \sum_{n=1}^{\infty} n |a_n|^2$$

as $k \to \infty$. Thus, we see that A is finite and (26) holds. The other assertions also follow in this way. We note incidentally that D always has area (Lebesgue measure) since D is either open or a point.

Once (26) is established and f satisfies $f'(0) = 1$ an immediate consequence is the inequality $A \geqq \pi$. Thus the map $z \to f(z)$ covers an area of at least π, and we see that $A = \pi$ only for the functions $f(z) = a_0 + z$.

We now consider the problem of minimizing $L(r,f)$ for functions analytic for $|z| < 1$ and satisfying $f'(0) = 1$. If a curve C is defined by the equation $w = w(\theta)$, where $a \leqq \theta \leqq b$, and $dw/d\theta$ is piecewise continuous then the length of that curve is given by $\int_a^b |w'(\theta)| \, d\theta$. This is consistent with the intuitive interpretation of $|w'(\theta)|$ as the local distortion of arc length given by the map $\theta \to w(\theta)$. The image of the circle $|z| = r$, where $0 < r < 1$, under f is given by the parametrization $w = f(re^{i\theta})$, where $0 \leqq \theta \leqq 2\pi$. Thus, $dw/d\theta = f'(re^{i\theta}) ire^{i\theta}$ and we thereby obtain the formula

$$L(r,f) = \int_0^{2\pi} |f'(re^{i\theta})| \, r \, d\theta. \tag{27}$$

Cauchy's formula applied to f' shows that

$$f'(0) = \frac{1}{2\pi i}\int_{|z|=r}\frac{f'(z)}{z}dz = \frac{1}{2\pi}\int_0^{2\pi} f'(re^{i\theta})\,d\theta. \tag{28}$$

The relations (28) and (27) imply that

$$|f'(0)| \leqq \frac{1}{2\pi}\int_0^{2\pi}|f'(re^{i\theta})|\,d\theta = \frac{1}{2\pi r}L(r,f). \tag{29}$$

Since $f'(0) = 1$ this is the same as

$$L(r,f) \geqq 2\pi r. \tag{30}$$

This is our geometric inequality about $L(r,f)$ and it asserts that the mapping $z \to f(z)$ transforms each circle $|z| = r$ onto a curve of length not less than the length of that circle. The "triangle inequality" applied to the above integral is an equality only when $f'(z)$ is constant, and thus $L(r,f) = 2\pi r$ only for the functions $f(z) = a_0 + z$.

Formula (27) and inequality (30) hold for $r = 1$ if suitably interpreted. For example, if f is continuous for $|z| \leqq 1$ and of bounded variation on $|z| = 1$, then

$$L = \int_0^{2\pi}|f'(e^{i\theta})|\,d\theta, \tag{31}$$

where the integral appropriately interpreted makes sense as a Lebesgue integral, and $L(r,f) \to L$ as $r \to 1$ [**45**, v. 1, p. 150]. This and (30) for $0 < r < 1$ implies that $L \geqq 2\pi$. The meaning of $f'(e^{i\theta})$ is given as a "boundary value function," that is, $f'(e^{i\theta})$ is defined by the limit

$$f'(e^{i\theta}) = \lim_{r\to 1} f'(re^{i\theta}) \tag{32}$$

which exists for almost all θ. L gives meaning to the notion of length of the curve traced out by f as z traces out $|z| = 1$, and it agrees with the usual idea of the length of such a curve when f is moderately smooth for $|z| \leqq 1$.

6. An improvement of $L(r,f) \geqq 2\pi r$. We shall show that the result expressed by inequality (30) can be refined by an application of the Principle of Subordination. Assume that f is analytic for $|z| < 1$, and as before, let $L(r,f)$ denote the length of the curve $w = f(re^{i\theta})$, $0 \leqq \theta \leqq 2\pi$, and let $D(r,f)$ denote the image of $|z| < r$ under f. Let $C(r,f)$ denote the boundary of the set $D(r,f)$ and let $\Gamma(r,f)$ denote the "outer boundary" of $D(r,f)$. We shall define outer boundary more precisely later but it is intuitively clear what it means. For example, if we take the disk $|z| < 1$ and punch out of it a finite number of closed disks, none of which intersect $|z| = 1$, then the resulting set has $|z| = 1$ as its outer boundary. The "inner boundary" of this set consists of af the perimeters of the disks punched out.

Let $l^*(r,f)$ denote the geometric length of the set $C(r,f)$ and let $l(r,f)$ denote the geometric length of $\Gamma(r,f)$. For example, if $f(z) = z^n$ where n is a positive integer then $l(r,f) = l^*(r,f) = 2\pi r^n$ as $D(r,f)$ is simply the set $|w| < r^n$. Recall also that in this example $L(r,f) = 2\pi n r^n$. Another example distinguishing these three lengths is illustrated by Figure 4, where a curve is shown representing the image of $|z| = r$ under f in the sense it is traced out once as z traces out $|z| = r$ once. In this example, $l(r,f)$ is the length of the curve $RPQR$, $l^*(r,f)$ is the sum of the lengths of the two curves $RPQR$ and $TVUT$, and $L(r,f)$ is the sum of the lengths of the three curves $RPQR$, $RSTWR$ and $TVUT$. In general, $L(r,f) \geqq l^*(r,f) \geqq l(r,f)$ and what we shall show is $l(r,f) \geqq 2\pi r$ given the normalization $f'(0) = 1$. In particular, this contains in it the assertion of (30) as well as the inequality $l^*(r,f) \geqq 2\pi r$ for the geometric length of the boundary of $D(r,f)$.

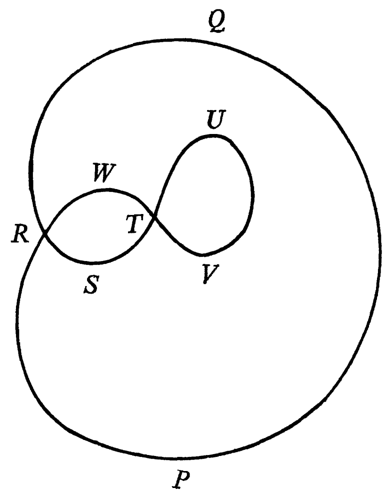


FIG. 4

First we indicate more precisely some of the geometric ideas involved. To each set D of complex numbers we associate another set H in the following way. Let E denote the complement of D in the extended plane, and let F be the component of E containing the point ∞. Let G be the complement of F in E and set $H = D \cup G$. Intuitively G consists of the "holes" of D, such as the punched out disks in our earlier example, and H may be viewed as the minimal set obtained from D which contains no holes. Therefore, the fact that H is simply connected is not surprising and we shall take advantage of this. (Recall that a set H is simply connected if it is open and connected and its complement is connected in the extended plane.) The **outer boundary** of D is defined to be the boundary of H.

The relation just described between D and H if applied to $D(r,f)$ yields a set denoted $H(r,f)$. If the set $H(r,f)$ could be associated with a univalent, analytic

function g in the sense that the image of $|z| < r$ under g is $H(r,f)$ and $g(0) = f(0)$, then we would have the exact relation that f is subordinate to g for $|z| < r$. At the same time the outer boundary of $D(r, f)$ becomes the boundary of $H(r,f)$.

The existence of such a function g is precisely what the Riemann Mapping Theorem asserts. We quote this theorem in the following form:

If H is any simply connected domain, different from the whole plane, and if $w_0 \in H$, then there is a function g that is analytic for $|z| < 1$ and maps $|z| < 1$ one-to-one onto H so that $g(0) = w_0$.

If the Riemann Mapping Theorem is applied to the set $H(r,f)$ where $w_0 = f(0)$, it produces a function g univalent and analytic for $|z| < 1$ and thus $f(rz)$ is subordinate to $g(z)$ for $|z| < 1$. Also, if L denotes the length of the curve $w = g(e^{i\theta})$, $0 \leqq \theta \leqq 2\pi$, then

$$L = l(r,f). \tag{33}$$

The relation (33) is intuitively clear, but to be more precise one needs to take advantage of the piecewise smoothness of $\Gamma(r,f)$ to show that g can be extended continuously for $|z| \leqq 1$ and is of bounded variation on $|z| = 1$. Also, as z traverses the circle $|z| = 1$ once, $g(z)$ traverses $\Gamma(r,f)$ once, always moving in the same direction.

If the result expressed by the inequality $L \geqq 2\pi$ is applied to the function $g(z)/g'(0)$, it implies that

$$L \geqq 2\pi |g'(0)|. \tag{34}$$

Since $h(z) = f(rz)$ is subordinate to $g(z)$ for $|z| < 1$ inequality (3) becomes $|h'(0)| \leqq |g'(0)|$, which is the same as $|g'(0)| \geqq r$ given the normalization $f'(0) = 1$. This combined with (33) and (34) produces our result, namely,

$$l(r,f) \geqq 2\pi r. \tag{35}$$

It is not difficult to see that equality in (35) occurs only for the functions $f(z) = a_0 + z$.

7. Translations of the range of an analytic function. The solution to a problem which we now discuss, depends in part on the way that subordination was used in the previous section. Specifically, subordination took place through the Riemann Mapping Theorem and the relation between the sets D and H. We shall also need to use some properties of the subordinating function, mainly that it is univalent. In particular, we eventually shall invoke the Koebe $\frac{1}{4}$-theorem. This problem has an added interest for our development here in that its solution (and formulation) depends on further interesting geometric relations.

If D is any set of complex numbers and b is a given complex number, then $D + b$ denotes the set of numbers of the form $d + b$, where $d \in D$. The set $D + b$ is a translation of D by the vector b, and $|b|$ is called the length of this translation.

Now let D denote the image of $|z| < 1$ under the analytic function f normalized

by $f'(0) = 1$. If f were additionally univalent for $|z| < 1$ and $f(0) = 0$, then D would contain the disk $|w| < \frac{1}{4}$, and thus every translation of D would meet D at least for translates of length less than $\frac{1}{2}$. The number $\frac{1}{2}$ is not the largest number for which this statement holds, but at least this asserts the existence of such a non-zero number. We shall determine the largest such number and find that it does not depend (directly) on the assumption that f is univalent. Specifically, we prove that if f is analytic for $|z| < 1$, $f'(0) = 1$, and D is the range of f, then each translation of D of length less than $\pi/2$ meets D. Moreover, $\pi/2$ is the largest number making this assertion true.

An outline of the proof of this is given. In part it depends on the geometric lemma: if $D + b \cap D = \varnothing$ then the sets $\{D + nb\}$, where n varies over the integers, are pairwise disjoint. A proof of this can be given depending on the concept of winding numbers. Its validity is intuitively suggested by Figure 5.

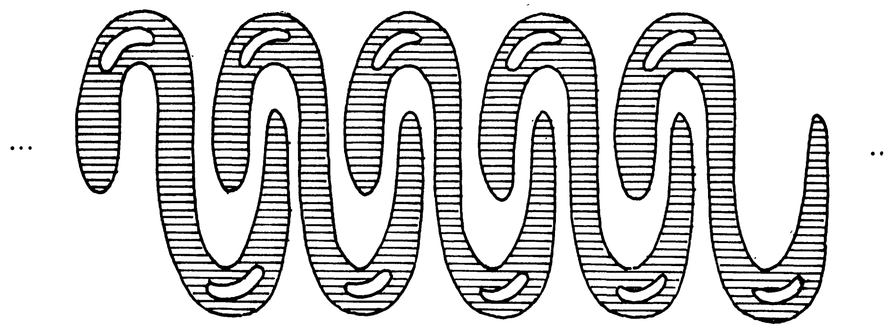


FIG. 5

Another geometric lemma is the fact that $D + b \cap D = \varnothing$ implies $H + b \cap H = \varnothing$. Here H is the set determined from D as discussed in section 6. From an intuitive viewpoint this lemma is quite expected, for if $D + b$ doesn't meet D, then the holes of $D + b$ cannot meet D nor can the holes of D meet $D + b$. The assertion is somewhat more general applying to domains similar to D; that is, if $|a| = 1$ and $aD + b \cap D = \varnothing$ then $aH + b \cap H = \varnothing$. A proof of this can be given using simple topological arguments. Such results presumably also hold for appropriate sets D in Euclidean n-space.

We now proceed to prove the above $\pi/2$-theorem. We suppose that $D + b \cap D = \varnothing$ and then need only show that $|b| \geqq \pi/2$. As $D + b \cap D = \varnothing$ we conclude that $H + b \cap H = \varnothing$, and in particular, H is not the whole plane. Since H is a simply connected domain containing the point $f(0)$, the Riemann Mapping Theorem implies the existence of an analytic, univalent function g which maps $|z| < 1$ onto H so that $g(0) = f(0)$. Thus, f is subordinate to g for $|z| < 1$ and accordingly $1 = |f'(0)| \leqq |g'(0)|$.

Since $H + b \cap H = \varnothing$ the sets $\{H + nb\}$, where n varies over the integers, are pairwise disjoint. This implies that the function $h(z) = (2\pi i/b)g(z)$ assumes no pair

of values differing by an integral multiple of $2\pi i$. Since $e^{z_1}=e^{z_2}$ implies that $z_2-z_1=2\pi n i$ for some integer n, and h is univalent for $|z|<1$, we conclude that $k(z)=e^{h(z)}$ also is univalent for $|z|<1$. The function $l(z)=(k(z)-k(0))/k'(0)$ is analytic and univalent for $|z|<1$ and satisfies $l(0)=0$ and $l'(0)=1$. Moreover, $l(z)\neq -k(0)/k'(0)$ for $|z|<1$ as k does not vanish. The Koebe $\frac{1}{4}$-theorem implies that $|-k(0)/k'(0)|\geqq \frac{1}{4}$, which is the same as $|b|\geqq(\pi/2)|g'(0)|$. As $|g'(0)|\geqq 1$, we conclude with the result of the theorem, namely $|b|\geqq\pi/2$.

The function

$$f(z)=\tfrac{1}{2}\log\frac{1+z}{1-z} \tag{36}$$

maps the disk $|z|<1$ one-to-one onto the domain $|\operatorname{Im} w|<\pi/4$ and satisfies $f'(0)=1$. These properties of f follow from the corresponding results for the mapping $w=(1+z)/(1-z)$. Also note that Im log $w=\arg w$ and log w is univalent on any set it is defined. This function shows that the $\pi/2$-theorem is "sharp," as its statement is no longer valid if $\pi/2$ is replaced by any larger number.

It is also interesting to note that the function

$$f(z)=\tfrac{1}{2}\log\frac{1+z}{1-z}$$

essentially serves as a subordinating function to prove the following: *if f is analytic for $|z|<1$, $f'(0)=1$ and D is the range of f, then the width of D in any direction is not less than $\pi/2$* [**36**, p. 130, problem 238]. This result is actually contained as a special case of the previous somewhat deeper $\pi/2$-theorem. In particular, we recall that if a set D has width $w(\theta)$ in the direction θ, then D is a subset of a strip with sides parallel to the vector $e^{i\theta}$ and such that those sides are $w(\theta)$ distance apart.

8. The principle of symmetrization. In this section we discuss the Principle of Symmetrization. It may be regarded as a relative of the Principle of Subordination. Each relates two sets corresponding to the ranges of two functions so that one of them has a larger value of $|f'(0)|$.

There are several kinds of symmetrization but we shall be interested only in Steiner symmetrization. In general, if D is a given domain then a symmetrization of D produces another domain D^* which has certain kinds of symmetry. The Steiner symmetrization of D with respect to a given line L is obtained by projecting the points of D toward L so that a set is obtained which is symmetric about L. More precisely, suppose that L is the real axis. The intersection of each line $x=a$ with D is a countable collection of open intervals having a total length $l(a)$, where $0\leqq l(a)\leqq\infty$. The Steiner symmetrization of D is the set

$$D^*=\{(x,y):|y|<\tfrac{1}{2}l(x)\}.$$

Since D is a domain its Steiner symmetrization D^* has a number of properties.

For example, it can be shown that D^* is a simply connected domain. Also the areas (Lebesgue measure) of D and D^* are equal.

If D^* is not the whole plane then by the Riemann Mapping Theorem it can be described as the range of a suitable univalent map of $|z| < 1$. Thus, when D is the range of a given function f the symmetrization of D will associate with f some univalent function g. The principle of symmetrization is the inequality $|f'(0)| \leqq |g'(0)|$. More precisely stated, let D be the range of the function f analytic for $|z| < 1$, and let D^* be the Steiner symmetrization of D with respect to any line through the point $f(0)$. If g is analytic and maps $|z| < 1$ one-to-one onto D^* such that $g(0) = f(0)$, then $|f'(0)| \leqq |g'(0)|$.

Although the statement of this principle of symmetrization is about as simple as that for subordination, its proof is exceedingly more difficult and lengthy. The applications of this principle to geometric function theory are quite varied and often elegant. We shall be content to give one such application relating to the inequality $A(r,f) \geqq \pi r^2$ discussed in section 5.

Let f be analytic for $|z| < 1$ and let $D(r,f)$ and $A(r,f)$ have their previous meaning. Also let $a(r,f)$ denote the area (Lebesgue measure) of the set $D(r,f)$. For example, if $f(z) = z^n$, where n is a positive integer, then $a(r,f) = \pi r^{2n}$ as $D(r,f)$ is simply the set $|w| < r^n$, whereas $A(r,f) = n\pi r^{2n}$. In the example represented by Figure 4, $a(r,f)$ is the sum of the areas of the two regions with boundaries $RSTWR$ and $RPQRWTUVTSR$, whereas $A(r,f)$ counts the area of the second region along with twice the area of the first region. In general, $A(r,f) \geqq a(r,f)$, and we shall show that $a(r,f) \geqq \pi r^2$, given that $f'(0) = 1$, thereby improving the earlier result $A(r,f) \geqq \pi r^2$. Thus, πr^2 is the least possible area of the set $D(r,f)$.

The proof is quite simple. The function $F(z) = f(rz)$ is analytic and maps $|z| < 1$ onto $D(r,f)$. Let D^* denote the Steiner symmetrization of $D(r,f)$ with respect to any line through $F(0) = f(0)$. As D^* is not the whole plane (it is even bounded) there is a function g analytic and univalent for $|z| < 1$ with the range D^* and such that $g(0) = F(0)$. By the principle of symmetrization $|F'(0)| \leqq |g'(0)|$, which is the same as $|g'(0)| \geqq r$ given that $f'(0) = 1$. The area of D^* satisfies $A \geqq \pi |g'(0)|^2$, because of the result given by the inequality $A \geqq \pi$ applied to the function $g(z)/g'(0)$. Since Steiner symmetrization preserves area, $A = a(r,f)$. Combining our results shows that $a(r,f) = A \geqq \pi |g'(0)|^2 \geqq \pi r^2$, that is,

$$a(r,f) \geqq \pi r^2. \tag{37}$$

This inequality, like (35), also holds when $r = 1$. For (37) this asserts that if f is analytic for $|z| < 1$ and $f'(0) = 1$ then the area of the range of f satisfies $A \geqq \pi$.

9. Comments on references and other results in geometric function theory. Our development here represents only some aspects of geometric function theory. We have limited the discussion to problems that have been of real concern to us and which can be presented simply without elaborate technical complications. This area

of mathematics has been influenced by some extraordinary mathematicians and has had a long and successful tradition. Today it remains a vital and active branch of mathematics.

Three excellent references for this paper are the books by Golusin [**11**], Hayman [**15**], and Nehari [**32**]. Various more general books on complex analysis contain information directly related to our development [for example, see **17**, v. 2, Chapters 17 and 18]. In [**4**] Bernardi presents a survey of the theory of univalent functions, and in [**5**] he has compiled an enormous list of books and research articles written up to 1966, concerning univalent functions. This bibliography is presented alphabetically by author and cross-referenced by topic, and is an excellent source for references. The papers [**26**] by Littlewood and [**40**] by Rogosinski contain the initial and basic results about subordination. The problem books [**36**] by Pólya and Szegö also contain results relating to several of our considerations.

The discussions of section 2 are contained in most standard books in complex analysis. A good reference to section 3 is [**32**, pp. 226ff]. The results of sections 4 and 5 can be found in [**11**], [**32**; see the problems on p. 155] and [**36**; see p. 140]. The development of sections 6, 7, and 8 are due to the author [see **27**, **28**]. A forthcoming paper [**29**] is based on ideas initiated in [**28**]. A good source for the principle of symmetrization is [**15**, Chapter 4].

There are various concepts or developments in complex analysis which can be appreciated if this paper has interested the reader. These are related to our presentation either through their geometric flavor or through analogous applications to analytic functions. Specifically, we mention the concepts of transfinite diameter or logarithmic capacity, harmonic measure and extremal length. Material on these ideas can be found in [**10**], [**11**], [**16**], and [**17**]. The study of quasiconformal mappings may also interest the reader [see **3**]. A more classical area is the study of the geometry of the zeros of a polynomial represented by the reference [**30**].

Distortion theorems like (8), (9), (10) and (11) have been of great concern to mathematicians for various families of analytic functions. For example, if S consists of the functions f analytic and univalent for $|z| < 1$ and normalized by $f(0) = 0$ and $f'(0) = 1$ there exist numerous such results. Specifically.

$$|f(z)| \leqq \frac{|z|}{(1-|z|)^2}$$

and this and some other distortion theorems can be obtained without much difficulty [see **15**, p. 4]. Other distortion theorems for S are more difficult to prove and depend on more elaborate analytic methods or more abstract considerations. For example, we mention the theory developed by Löwner [**15**, Chapter 6] and the methods represented by the books [**18**] and [**41**].

In section 4 we mentioned that the condition $\operatorname{Re}\{zf'(z)/f(z)\} > 0$ (and $f'(0) \neq 0$) for $|z| < 1$ is equivalent to the condition that f be (univalent and) starlike for $|z<1|$.

Similarly, if f is analytic for $|z| < 1$ and $f'(0) \neq 0$, then the condition

$$\operatorname{Re}\left\{\frac{zf''(z)}{f'(z)} + 1\right\} > 0 \text{ for } |z| < 1$$

is necessary and sufficient for f to map $|z| < 1$ one-to-one onto a convex domain. There are various other conditions on a function sufficient to imply that it be univalent or that its range have a certain geometric property. Some of these are mentioned in [**4**]. For example, if f is analytic in a convex domain D and if $\operatorname{Re} f'(z) > 0$ for all z in D, then f is univalent in D [**22**, p. 582]. These considerations are much more varied, but they do have some similarity to the implications given by $f'(x) > 0$ or $f''(x) > 0$ (for $a < x < b$) in elementary calculus.

There are many additional results about the geometry of the range D of a function analytic for $|z| < 1$ and satisfying $f'(0) = 1$. For example, the diameter of D is not less than 2 [**25**; **36**, p. 130, problem 239]. Similar results about the diameter of the range of a function are found in [**7**] and in [**8**] where the function is analytic in an annulus.

Furthermore, if f is univalent for $|z| < 1$ and $f(0) = 0$ so that $f \in S$, then not only does D contain the disk $|w| < \frac{1}{4}$ but each such set D contains some open disk of radius R with $R > \frac{1}{4}$. The largest value of R satisfying this last assertion is called the Bloch-Landau constant ("for univalent functions"). The exact value of R is unknown although several contributions have been made in this direction [see **1**, **2**, **14**, **20**, **24**, **37**, **43**]. If D is convex then D will contain some open disk of radius $\pi/2$, and $\pi/2$ is the largest number with this property [**44**]. Let A_f denote the area of the intersection of D with the disk $|w| < 1$ and let $A = \inf A_f$, where f varies over S. The problem of finding A was raised in [**12**] and although some results were obtained there and later improved in [**19**] and [**13**] the exact value of this "fixed area" is still unknown. Another interesting geometric result is the following: there is a line segment in D with one endpoint at $w = 0$ and with a length greater than 0.73 [**42**]. The largest number that 0.73 can be replaced by so that this statement holds is still an open problem.

Several interesting "covering theorems" are obtained in [**32**], [**33**], [**34**], and [**35**] as, for example, the following: *Let f be analytic for $|z| < 1, f(0) = 0, f'(0) = 1$ and let $f(z) \neq 0$ for $z \neq 0$, $|z| < 1$; then the range of f contains the disk $|w| < 1/16$* [**32**, p. 323]. The results we refer to are obtained by subordination and the subordinating function is defined in terms of the elliptic modular function. This represents a much more difficult situation than that discussed in this paper since our subordinating function was often a very simple function. The papers quoted above also use subordination in a more general sense, where f^{-1} is only locally defined and analytic, but nevertheless, $\phi = f^{-1}(g)$ becomes a well-defined analytic, function satisfying Schwarz's lemma. Perhaps the most notable success of the use of subordination in this context and associated with the elliptic modular function is the famous

Picard theorem: *Each non-constant entire function takes on every complex number with at most one exception* [**32**, p. 321].

We have discussed the problem of finding the minimum of the quantities $L(r,f)$ and $A(r,f)$ given that f is analytic for $|z| < 1$ and $f'(0) = 1$. The problem of maximizing these two quantities has been considered by various mathematicians for a number of families of functions f. When f is in S specific upper bounds for $L(r,f)$ and $A(r,f)$ are known but the precise upper bounds have not yet been determined. For the subfamilies of S consisting of convex, or starlike or "close-to-convex" [see **21**] functions the best upper bounds for $L(r,f)$ and $A(r,f)$ are known [**6**, **23**, **31**]. Similar results hold for the class of "typically real functions" or for functions having a positive real part for $|z| < 1$ (see [**39**] for some of these results).

We finally mention a theorem due to Fejér and Riesz [**9**]. *If g is analytic for $|z| < 1$ and continuous for $|z| \leqq 1$ then*

$$\int_{-1}^{1} |g(re^{i\theta})|\, dr \leqq \frac{1}{2} \int_{-\pi}^{\pi} |g(e^{i\phi})|\, d\phi . \tag{38}$$

If we set $g = f'$ where f is analytic for $|z| < 1$ then this inequality asserts that $\lambda \leqq \frac{1}{2}L$, where λ is the length of the image of the diameter $-1 \leqq x \leqq 1$ under f, and L is the length of the image of $|z| = 1$ under f. Both λ and L count lengths as given by parametrization through f. Also, the number $\frac{1}{2}$ in $\lambda \leqq \frac{1}{2}L$ cannot in general be replaced by a smaller number.

A more complete listing of appropriate references can be obtained by consulting [**5**] in the topic reference list. In particular, references are given there under the headings "the principle of subordination," "relations involving arc length," "relations involving area," and "covering theorems."

This paper is an expanded version of a talk given at the meeting of The Mathematical Association of America held in Williamstown, Massachusetts, on June 21, 1969. Earlier similar ideas were presented at student-faculty colloquia both at Lafayette College and at Williams College.

References

1. L. V. Ahlfors and H. Grunsky, Über die Blochsche Konstante, Math. Z., 42 (1937) 671–673.

2. L. V. Ahlfors, An extension of Schwarz's lemma, Trans. Amer. Math. Soc., 43 (1938) 359–364.

3. ———, Lectures on Quasiconformal Mappings, Van Nostrand, Princeton, N. J. 1966.

4. S. D. Bernardi, A survey of the development of the theory of schlicht functions, Duke Math. J. 19 (1952) 263–287.

5. ———, Bibliography of Schlicht Functions, Courant Institute of Mathematical Sciences (New York University), New York, 1966.

6. P. L. Duren, An arclength problem for close-to-convex functions, J. London Math. Soc., 39 (1964) 757–761.

7. I. Dziubinski, Sur le minimum du diamètre de la fonction transformant la région en cercle unité, Bull. Soc. Sc. Lett. Lodz, 11 (1960) no. 3, 3 pp.

8. ———, Sur le minimum de diamètre du domaine doublement convexe, Bull. Soc. Sc. Lett. Lodz, 14 (1963) no. 6, 9 pp.

9. L. Fejér and F. Riesz, Über einige funktionentheoretische Ungleichungen, Math. Z., 11 (1921) 305–314.

10. W. H. J. Fuchs, Topics in the Theory of Functions of one Complex Variable, Van Nostrand, Princeton, N. J., 1967.

11. G. M. Goluzin, Geometric Theory of Functions of a Complex Variable, Amer. Math. Soc. 1969.

12. A. W. Goodman, Note on regions omitted by univalent functions, Bull. Amer. Math. Soc., 55 (1949) 363–369.

13. ———, and E. Reich, On regions omitted by univalent functions II, Canadian J. Math., 7 (1955) 83–88.

14. R. E. Goodman, On the Bloch-Landau constant for schlicht functions, Bull. Amer. Math. Soc., 51 (1945) 234–239.

15. W. K. Hayman, Multivalent Functions, Cambridge University Press, Cambridge, 1958.

16. M. Heins, Selected Topics in the Classical Theory of Functions of a Complex Variable, Holt, Rinehart and Winston, New York, 1962.

17. E. Hille, Analytic Function Theory, Ginn, Boston, 1962.

18. J. A. Jenkins, Univalent Functions and Conformal Mapping, Springer, Berlin, 1958.

19. ———, On values omitted by univalent functions, Amer. J. Math., 75 (1953) 406–408.

20. ———, On the schlicht Bloch constant, J. Math. Mech., 10 (1961) 729–734.

21. W. Kaplan, Close-to-convex schlicht functions, Michigan Math. J., 1 (1952) 169–185.

22. ———, Advanced Calculus, Addison-Wesley, Reading, Mass., 1952.

23. F. R. Keogh, Some inequalities for convex and starshaped domains, J. London Math. Soc., 29 (1954) 121–123.

24. E. Landau, Über die Blochsche Konstante und zwei verwandte Weltkonstanten, Math. Z., 30 (1929) 608–634.

25. ———, and O. Toeplitz, Concerning the greatest oscillation of an analytic function in a circle, Archiv Math. Physik, (3) 11 (1907) 302.

26. J. E. Littlewood, On inequalities in the theory of functions, Proc. London Math. Soc., (2) 23 (1925) 481–519.

27. T. H. MacGregor, Length and area estimates for analytic functions, Michigan Math. J. 11 (1964) 317–320.

28. ———, Translations of the image domains of analytic functions, Proc. Amer. Math. Soc., 16 (1965) 1280–1286.

29. ———, Rotations of the range of an analytic function (to appear).

30. M. Marden, Geometry of Polynomials, Amer. Math. Soc., 1966.

31. A. Marx, Untersuchungen über schlichte Abbildungen, Math. Ann., 107 (1932–1933) 40–67.

32. Z. Nehari, Conformal Mapping, McGraw-Hill, New York, 1952.

33. ———, On analytic functions possessing certain properties of univalency, Proc. London Math. Soc., (2) 50 (1945) 120–136.

34. ———, The elliptic modular function and a class of analytic functions first considered by Hurwitz, Amer. J. Math., 69 (1947) 70–86.

35. ———, A generalization of Schwarz's lemma, Duke Math. J., 14 (1947) 1035–1049.

36. G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, Springer, Berlin, 1954.

37. E. Reich, On a Bloch-Landau constant, Proc. Amer. Math. Soc., 7 (1956) 75–76.

38. R. M. Robinson, The Bloch constant A for a schlicht function, Bull. Amer. Math. Soc., 41 (1935) 535–540.

39. W. Rogosinski, Über positive harmonische Entwicklungen und typisch-reelle Potenzreihen, Math Z., 35 (1932) 93–121.

40. ———, On the coefficients of subordinate functions, Proc. London Math. Soc., (2) 48 (1943) 48–82.

41. A. C. Schaeffer and D. C. Spencer, Coefficient Regions for Schlicht Functions, Amer. Math. Soc., 1950.

42. E. Strohhäcker, Beiträge zur Theorie der schlichten Funktionen, Math. Z., 37 (1933) 356–380.

43. C. Ulucay, Bloch functions of the third kind and the constant A, Proc. Amer. Math. Soc., 8 (1957) 923–925.

44. M. Zhang, Ein Überdeckungssatz für konvexe Gebiete, Acad. Sinica Sci. Rec., 5 (1952) 17–21.

45. A. Zygmund, Trigonometric Series, Cambridge University Press, Cambridge, 1959.

HORIZONTAL CHORD THEOREMS

J. C. OXTOBY, Bryn Mawr College

For any real function f, defined on a bounded or unbounded interval, the set

$$H(f) = \{h \in [0,\infty): f(x) = f(x+h) \text{ for some } x\}$$

is called the **chord set** of f. The purpose of this note is to present some generalizations of known theorems concerning these sets, and to test their generality by means of counter examples.

1. Functions having every chord. It is well known, and very easy to prove [**1**, p. 78], that if f is periodic and continuous on the real line R, then $H(f)=[0,\infty)$; briefly, a continuous periodic function has every chord. This result was generalized by Diaz and Metcalf [**3**], who showed that it is sufficient to assume that f is periodic, continuous at some point, and that for each $h>0$, the function $f(x+h)-f(x)$ has a connected range. In particular, a periodic derivative has every chord. It is not sufficient to assume that f itself has a connected range, or even to assume that f is a Darboux function of Baire class 1. (A function f is a Darboux function if its domain is an interval and if it has the intermediate value property, that is, maps each subinterval onto a connected set. It is of Baire class 1 if it can be represented as the limit of a convergent sequence of continuous functions.) For example,

$$f(x) = \cos\left(\frac{1}{\sin x}\right) + \tfrac{1}{2}(-1)^{[x/\pi]}, \quad \cos\left(\frac{1}{0}\right) = 0, \tag{1}$$

where $[x]$ denotes the largest integer less than or equal to x, is a Darboux function of Baire class 1 with period 2π, but it has no chord of length π.

* Professor Oxtoby earned his A. B. and M. A. degrees at Berkeley, and he was a Junior Fellow of the Harvard Society of Fellows. Except for a year's leave at Yale University, he has been at Bryn Mawr since he finished his graduate work. He was the Hedrick Lecturer in 1956, and he is the author of *Measure and Category* (Springer, 1971). His main research interests are measure and ergodic theory. *Editor.*

The hypothesis that f be continuous at some point cannot be omitted. Let B be a Hamel basis (a maximal set of real numbers linearly independent over the rationals) that includes $b_1 = 1$. Define $f(\sum_1^n x_i b_i) = \sum_2^n x_i b_i$ whenever $b_1, \cdots, b_n$ are distinct members of B and $x_1, \cdots, x_n$ are rational numbers. Then $f(x+h) - f(x) = f(h)$ for all x and h. This function is periodic (every rational number is a period), and each of the functions $f(x+h) - f(x)$ is continuous (in fact, constant), but f has no chord of irrational length.

Generalizing in another direction, Tews [7] showed that a continuous *almost* periodic function has every chord. Actually, a much simpler and more general theorem holds, as we shall now show.

Let us say that a function f, defined on an interval, is **positively recurrent at** x_0 if for every $\varepsilon > 0$ the set $\{x : |f(x) - f(x_0)| < \varepsilon\}$ is unbounded above, and that f is **positively recurrent** if it is positively recurrent at each point of its domain. Note that the domain of such a function must be of the form (a, ∞), $[a, \infty)$, or $(-\infty, \infty)$. Replacing "above" by "below" gives the corresponding definitions for **negatively recurrent**. These definitions are consistent with the notion of recurrence used in topological dynamics [4]. Since an almost periodic function is obviously recurrent (both positively and negatively), Tews's result is a corollary of the following theorem.

THEOREM 1. *If f is continuous and either positively or negatively recurrent on an interval, then f has every chord.*

Proof. Suppose some positive number h does not belong to $H(f)$. Then $f(x+h) - f(x)$ never changes sign. Since $f(x)$, $f(-x)$, $-f(x)$, and $f(x-c)$ all have the same chord set, we may assume that f is positively recurrent, that $f(x+h) > f(x)$ for all x, and that the domain of f includes $[0, \infty)$. Let f attain its minimum value on $[0, h]$ at x_0, and its minimum on $[h, 2h]$ at x_1. If $x > h$, then $x - nh \in (h, 2h]$ for some integer $n \geqq 0$, and

$$f(x) \geqq f(x - nh) \geqq f(x_1) > f(x_1 - h) \geqq f(x_0).$$

Thus, when $\varepsilon = f(x_1) - f(x_0)$, the set

$$\{x : |f(x) - f(x_0)| < \varepsilon\}$$

is bounded above by h. This contradicts the hypothesis that f is positively recurrent.

In Theorem 1 it is not sufficient to assume that f is positively or negatively recurrent at each point. Any horizontal line that meets the graph of

$$f(x) = 2\sin 2\pi x + \tanh x \tag{2}$$

meets it in an unbounded set. (When $|y| < 1$ the set $f^{-1}(y)$ is even relatively dense.) But f has no chord of length 1. More surprising is the fact that a recurrent derivative need not have every chord. For each x in R, let m be the largest integer less than x and define

$$f(x) = 2^m\left(1 + x - m + \sin\frac{\pi}{x - m}\right). \tag{3}$$

This function is continuous when x is not an integer. When n is an integer, f is continuous on the left but assumes all values between 0 and 2^{n+1} in every right neighborhood of n. It follows that the set $\{x : f(x) = f(x_0)\}$ is unbounded above, for each x_0. Hence f is positively recurrent. To verify that f is the derivative of the function

$$F(x) = \int_{-\infty}^{x} f(t)\,dt,$$

note that when n is an integer and $0 < h < 1$ we have

$$F(n + h) - F(n) - hf(n) = 2^{n-1}h^2 + 2^n \int_0^h \sin\frac{\pi}{t}\,dt.$$

Replacing $\sin \pi/t$ by

$$\frac{d}{dt}\left(\frac{t^2}{\pi}\cos\frac{\pi}{t}\right) - \frac{2t}{\pi}\cos\frac{\pi}{t},$$

the right member takes the form

$$2^{n-1}h^2 + \frac{2^n h^2}{\pi}\cos\frac{\pi}{h} - \frac{2^{n+1}}{\pi}\int_0^h t\cos\frac{\pi}{t}\,dt,$$

which is numerically less than $2^{n+1}h^2$. Hence F has a right derivative at n equal to $f(n)$. The fundamental theorem then completes the proof that $F'(x) = f(x)$ everywhere. Nevertheless, f has no chord of length 1, since $f(x + 1) - f(x) = f(x) > 0$ for all x.

2. The universal chord theorem. For continuous functions, the universal chord theorem of P. Lévy [**6**] [see also **1**, p. 79] asserts:

(i) *If $h \in H(f)$, then $h/n \in H(f)$ for every positive integer n.*

(ii) *If a and h are positive numbers and a is not a submultiple of h, then there exists a continuous function f with $h \in H(f)$ and $a \notin H(f)$.*

Lévy's example for (ii) was

$$f(x) = \sin^2\left(\frac{\pi x}{a}\right) - \frac{x}{h}\sin^2\left(\frac{\pi h}{a}\right). \tag{4}$$

The proof of (i) depends only on the fact that each of the functions $f[x + (h/n)] - f(x)$ has the intermediate value property. Hence the theorem holds for derivatives, as Boas [**1**, p. 81] has remarked, and also for approximate derivatives [**2**, p. 31]. However, the function defined by equation (1) shows that the theorem can fail for

a Darboux function of Baire class 1, despite the fact that any such function is topologically equivalent to a derivative, by Maximoff's theorem [**2**, p. 49].

3. A stronger form of Hopf's theorem. H. Hopf [**5**] obtained a complete characterization of the chord sets of functions continuous on a bounded closed interval and of plane continua. (If K is a subset of the plane, then

$$\{h \in [0, \infty): (x + h, y) \in K \text{ for some } (x, y) \in K\}$$

is called the **chord set** of K.) A subset of $(0, \infty)$ is called **additive** if it contains the sum of any two of its members. Let us call a set $H \subset [0, \infty)$ **co-additive** if $0 \in H$ and the set $H^* = (0, \infty) - H$ is additive. Hopf's theorem reads as follows:

(i) *The chord set of any non-empty compact connected subset of the plane is compact and co-additive.*

(ii) *Any compact co-additive subset of* $[0, \infty)$ *is the chord set of some plane continuum; more particularly, it is the chord set of some continuous function on a bounded closed interval.*

The function that Hopf used to prove (ii) was not differentiable. Nevertheless, the following theorem is true.

THEOREM 2. *For any compact co-additive set* $H \subset [0, \infty)$ *there exists a function* F *of class* C^∞ *on* R *such that* H *is the chord set of* F *and also of the restriction of* F *to* $[0, B]$, *where* $B = \sup H$.

Proof. We shall obtain this result by smoothing Hopf's function. Accordingly, we begin by repeating his proof of (ii), with a few minor changes. Let $\dot{H}$ denote the boundary of H, and define

$$f(x) = \begin{cases} d(x, \dot{H}) & \text{for } x \in H \cup (-\infty, 0) \\ -d(x, \dot{H}) & \text{for } x \in H^*, \end{cases} \tag{5}$$

where d denotes ordinary distance in R.

Note that 0 and B belong to $\dot{H}$, and that H^* is open. If we regard H as a subset of the x-axis, the graph of f has a right-angled peak, with endpoints in $\dot{H}$, above each component of $H - \dot{H}$, and a right-angled trough, with endpoints in $\dot{H}$, below each component of $(0, B) - H$. It also includes the points of $\dot{H}$ itself and the rays with slope -1 to the left of the origin and to the right of B.

If $B = 0$, then $f(x) = -x$ for all x, and the theorem is true in this case. Hence we may assume $B > 0$. Clearly f is continuous and satisfies a Lipschitz condition. It is not differentiable, but $f'(x) = \pm 1$ for any x that is not in $\dot{H}$ and is not the midpoint of one of the components of $(0, B) - \dot{H}$. The proof that H is the chord set of f and of its restriction f_B to $[0, B]$ rests on two lemmas:

1° *If* H *contains an interval of length* λ, *then* $(0, \lambda) \subset H$.

2° *If* $a \in \dot{H} \cup H^*$ *and* $a \leqq a + h \in H$, *then* $h \in H$.

Statement 1° follows from the fact that if $0 < a < \lambda$, then any interval of length λ contains a multiple of a. Consequently, no element of H^* can belong to $(0, \lambda)$. To prove 2°, let the hypotheses be satisfied and suppose $h \notin H$. Then h belongs to H^*, and therefore $(h-\varepsilon, h+\varepsilon) \subset H^*$ for some $\varepsilon > 0$. Moreover, a cannot be 0. Hence a belongs to $\dot{H} - \{0\}$ or to H^*. In either case, a is a cluster point of H^*. Therefore $a + x \in H^*$ for some $|x| < \varepsilon$, and $h - x \in H^*$. By additivity, $(a + x) + (h - x) = a + h$ belongs to H^*, contrary to hypothesis.

If $h \in \dot{H}$, then $f(0) = f(h) = 0$. Since $0 \leqq h \leqq B$, it follows that $h \in H(f_B)$. If $h \in H - \dot{H}$, let $h + 2a$ be the first point of $\dot{H}$ to the right of h. (Such a point exists, since $h < B \in \dot{H}$.) Then $h + 2a$ is a point of $\dot{H}$ nearest to $a + h$, and $f(a+h) = a > 0$. Since $(h, h + 2a) \subset H$, 1° implies that $(0, 2a) \subset H$. Consequently, 0 is a point of $\dot{H}$ nearest to a, and therefore $f(a) = a = f(a + h)$. Since $0 < a < a + h < B$, it follows that $h \in H(f_B)$. Thus $H \subset H(f_B)$.

If $h \in H(f)$, then $h \geqq 0$ and $f(x) = f(x + h)$ for some $x \in R$. To show that $h \in H$ we distinguish three cases. If $f(x) = f(x + h) = 0$, then both x and $x + h$ belong to $\dot{H}$, and the conclusion follows from 2°, with $a = x$. If $f(x) = f(x + h) > 0$, let a be a point of $\dot{H}$ nearest to x. Since $f(x + h) = |x - a|$, the interval with endpoints $x + h \pm |x - a|$ is contained in H. Its endpoints also belong to H; in particular, $a + h \in H$. Then 2° implies that $h \in H$. Lastly, if $f(x) = f(x + h) < 0$, let $a + h$ be a point of $\dot{H}$ nearest to $x + h$. Then $f(x) = -|x-a|$. All points of the open interval with endpoints $x \pm |x - a|$ belong to H^*. One of these endpoints is a. Hence a belongs either to H^* or to $\dot{H}$, and 2° implies that $h \in H$. This completes the proof that $H = H(f) = H(f_B)$.

(The foregoing argument becomes intuitively clear if one observes that when a chord of f lying above the x-axis is slid downward, keeping its left endpoint on the graph of f, its right endpoint ends up in a point of H; and when a chord lying below the x-axis is slid upward, keeping its right endpoint on the graph of f, its left endpoint ends up in a point of H^* or in a point of $\dot{H}$.)

To obtain Theorem 2 from this result, observe that if ϕ is any 1-1 mapping of R into R, then the composite function $F = \phi \circ f$ has the same chord set as f. By suitable choice of ϕ, we shall show that F can be made to be of class C^∞.

The open set $E = (0, B) - \dot{H}$ has only a finite number of components whose length exceeds any given positive number. Let $y_0 > y_1 > y_2 > \cdots$ be a strictly decreasing sequence of positive numbers, tending to 0, such that the half-length of each component of E is a term of the sequence. Before defining ϕ, we shall show that if ϕ is of class C^∞ on R, and if

$$\phi^{(n)}(0) = \phi^{(n)}(y_i) = \phi^{(n)}(-y_i) = 0 \tag{6}$$

for $n \geqq 1$ and $i \geqq 0$, then $F = \phi \circ f$ is of class C^∞ on R.

Let I be any component of E, and denote its midpoint by x_1. Then $f(x_1) = \pm y_i$ for some $i \geqq 0$. On one half of I we have $f' = 1$; on the other half, f' is equal to -1. On both of these intervals, $|F^{(n)}| = |\phi^{(n)} \circ f|$ for $n \geqq 1$, and (6) implies that

$F^{(n)}(x) \to 0$ as $x \to x_1$. Since F is continuous, it follows by induction and l'Hospital's Rule that $F^{(n)}(x_1) = 0$ for $n \geqq 1$. Thus $|F^{(n)}| = |\phi^{(n)} \circ f|$ on each component of E, and also on $(-\infty, 0)$ and (B, ∞), for $n \geqq 0$.

Let $x_0 \in \dot{H}$. If $x \notin \dot{H}$, the component of $R - \dot{H}$ to which x belongs has an endpoint a in $[x_0, x)$ or in $(x, x_0]$, and $a \in \dot{H}$. By the mean value theorem,

$$|F(x) - F(x_0)| = |F(x) - F(a)| = |(x-a)F'(\xi)| \leqq |x - x_0| \cdot |\phi'[f(\xi)]|$$

for some ξ between a and x. It follows from (6) and the constancy of F on $\dot{H}$ that $F'(x_0) = 0$. Assuming $F^{(n)} = 0$ on $\dot{H}$, similar reasoning shows that $F^{(n+1)} = 0$ on $\dot{H}$. Thus, by induction, F is of class C^∞ on R, and all its derivatives vanish on $\dot{H}$ as well as at the midpoints of the components of E.

It only remains to define a function ϕ having all the properties we have assumed. Recall that the function

$$\omega(x) = \exp\left(-\frac{1}{x^2}\right), \quad \omega(0) = 0, \tag{7}$$

is of class C^∞ on R and that $\omega^{(n)}(0) = 0$ for $n \geqq 0$. For each positive integer i, let b_i be an upper bound of the function $\omega(x - y_i) \cdot \omega(x - y_{i-1})$ and of the absolute values of its first $i - 1$ derivatives on the interval $[y_i, y_{i-1}]$. Put $a_i = 1/ib_i$ and define

$$\psi(x) = \begin{cases} \omega(x - y_0) & \text{on } (y_0, \infty) \\ a_i\omega(x - y_i) \cdot \omega(x - y_{i-1}) & \text{on } (y_i, y_{i-1}] \end{cases} \tag{8}$$

for $i \geqq 1$. Then define $\psi(0) = 0$ and $\psi(x) = \psi(-x)$ when $x < 0$. It is clear that ψ and each of its derivatives tends to 0 at each of the points y_i, and also as $x \to 0$, since

$$|\psi^{(n)}(x)| \leqq 1/i \text{ on } [y_i, y_{i-1}] \quad \text{for } i > n \geqq 0.$$

Therefore, ψ is of class C^∞ on R. Moreover, $\psi(x) > 0$ except at the points $0, \pm y_0, \pm y_1, \cdots$, where it vanishes together with its derivatives of all orders. Consequently, the function

$$\phi(x) = \int_0^x \psi(t)\, dt \tag{9}$$

is strictly increasing, of class C^∞ on R, and satisfies conditions (6). This completes the proof of Theorem 2.

4. Chord sets of analytic functions. Is it possible to find an *analytic* function having any prescribed chord set? (Recall that Lévy used an elementary function (4) to satisfy the much less stringent requirements of part (ii) of his theorem.) To see that no such improvement of Theorem 2 is possible, let C be a nowhere dense perfect

subset of $[1,2]$, and take $H = [0,1] \cup C$. Then H is compact, $0 \in H$, and $H^* = (1, \infty) - C$ is additive. Since $\dot{H} = \{0,1\} \cup C$ is uncountable, the following theorem shows that no analytic function can have H for its chord set.

THEOREM 3. *If $f:[a,b] \to R$ is continuous on $[a,b]$ and analytic on (a,b), then the boundary $\dot{H}$ of $H = H(f)$ is countable.*

Proof. If f is constant, then $\dot{H} = \{0, b-a\}$ has only two elements. We may therefore assume that f is not constant on (a,b). Then the set

$$D = \{x \in (a,b): f'(x) = 0\} \cup \{a,b\}$$

is countable, since the zeros of f' are isolated. The set

$$E_1 = \{h \in [0, b-a]: f(x) = f(x \pm h) \text{ for some } x \in D\}$$

is also countable, since f can assume any value at most countably many times. Let E_2 denote the set of endpoints of components of the open set $H - \dot{H}$. Evidently E_2 is countable. We shall show that $\dot{H} - (E_1 \cup E_2)$ is finite.

Let $h \in \dot{H} - (E_1 \cup E_2)$. Then $h \in H$ and $f(x_0) = f(x_0 + h)$ for some $x_0 \in [a, b-h]$. Since $h \notin E_1$, neither x_0 nor $x_0 + h$ can belong to D, hence

$$a < x_0 < x_0 + h < b,\ f'(x_0) \neq 0, \text{ and } f'(x_0 + h) \neq 0.$$

It follows that f is locally invertible at x_0 and at $x_0 + h$; there exist continuous functions ϕ and ψ, defined on an open interval I containing $y_0 = f(x_0)$, such that $\phi(y_0) = x_0$, $\psi(y_0) = x_0 + h$, and

$$f[\phi(y)] = f[\psi(y)] = y \text{ for all } y \in I. \tag{10}$$

Since $\psi(y_0) - \phi(y_0) = h > 0$, we may assume that $\psi(y) - \phi(y) > 0$ on I. Then (10) implies that $\psi(y) - \phi(y) \in H$ for all $y \in I$. Since $\psi - \phi$ is continuous, it maps I onto a connected subset of H that contains h. Since $h \in \dot{H} - E_2$, no connected subset of H can contain h and a number different from h. Therefore $\psi(y) - \phi(y) = h$ for all $y \in I$. Putting $x = \phi(y)$, it follows that $f(x) = f(x+h)$ on the subinterval $\phi(I)$ of (a,b). By analytic continuation, this equation holds for all $a < x < b + h$. Hence h is a member of the set

$$E_3 = \{h \in (0, b-a): f(x) = f(x+h) \text{ on } (a, b-h)\}.$$

Thus $\dot{H} - (E_1 \cup E_2) \subset E_3$.

If h_1 and h_2 are in E_3 and $h_1 < h_2$, then

$$a < a + h_1 < b - h_2 + h_1 < b.$$

If x is in the interval $J = (a + h_1, b - h_2 + h_1)$, then $x - h_1$ is in both $(a, b - h_2)$ and $(a, b - h_1)$. From the definition of E_3 it follows that $f(x - h_1) = f(x - h_1 + h_2)$ and $f(x - h_1) = f(x - h_1 + h_1) = f(x)$. Therefore $f(x) = f(x + h_2 - h_1)$ on J.

By analytic continuation, this equation holds on $(a, b - h_2 + h_1)$, hence $h_2 = h_1 \in E_3$. Thus E_3 contains positive differences of its members. If E_3 were infinite it would follow that E_3 has arbitrarily small members, and then f would be constant. Therefore E_3 must be finite. Consequently, $\dot{H}$ is countable.

References

1. R. P. Boas, Jr., A Primer of Real Functions, Carus Mathematical Monograph No. 13, Wiley, New York, 1960.

2. A. M. Bruckner and J. L. Leonard, Derivatives, this MONTHLY, 73 (April 1966) Herbert Ellsworth Slaught Memorial Papers, No. 11, 24–56.

3. J. B. Diaz and F. T. Metcalf, A continuous periodic function has every chord twice, this MONTHLY, 74 (1967) 833–835.

4. W. H. Gottschalk and G. A. Hedlund, Recursive properties of transformation groups, Bull. Amer. Math. Soc., 52 (1946) 637–641.

5. H. Hopf, Über die Sehnen ebener Kontinuen und die Schleifen geschlossener Wege, Comment. Math. Helv., 9 (1937) 303–319.

6. P. Lévy, Sur une généralisation du théorème de Rolle, C. R. Acad. Sci. Paris, 198 (1934) 424–425.

7. M. C. Tews, A continuous almost periodic function has every chord, this MONTHLY, 77 (1970) 729–731.

WOMEN IN MATHEMATICS

MARY GRAY, The American University

As I looked out over the audience at the Monday afternoon session of the 1971 Summer MAA meeting, I observed that over twenty percent were women, far more than at any other MAA session in my memory. This influx of female mathematicians was due, of course, to the subject of the panel: Women in Mathematics. Indeed, many in the audience told me later that they had come to the Penn State meeting solely or primarily because of the panel.

Under the direction of its moderator, Christine Ayoub of Penn State, the panel decided to focus on two questions: 1. Is there discrimination against women in mathematics? 2. What can, or should, be done to improve the status of women in the field?

The positioning of the members came to be symbolic, with the conservatives — the moderator and panelist Mary Ellen Rudin of Wisconsin on the right and the more militant Gloria Hewitt of Montana and myself on the left. There was an omission — there probably should have been representation of graduate students and/or assistant professors on the panel rather than only those who have more or less "made it."

All panelists agreed on the paucity of women in mathematics and that the condition becomes more pronounced as the level rises — high school, college, beginning years of graduate work, Ph.D. level, faculty positions; there are plenty of statistics to back up this impression. Many reasons were advanced to account for the dropout rate of women: cultural conditioning, inability to think abstractly, lack of commitment to the concentrated effort required by mathematical research, family pressures, etc. Rudin in particular held that there is little if any overt discrimination; other panelists and some audience members disagreed, maintaining that the statistics themselves are *prima facie* evidence of discrimination. For example, only six percent of the Ph.D.'s awarded in mathematics in recent years went to women although nearly half of the freshmen in mathematics classes are women. Even worse, no Sloan fellowship has ever gone to a woman in pure mathematics (currently seventy fellowships are awarded yearly in physical sciences and mathematics and the program is twenty years old). The faculties of the schools rated as the top twenty-seven in the 1969 survey of graduate schools show only a handful of women in tenured positions. When challenged to list great women mathematicians few are able to get farther than E. Noether, S. Kovalevsky, and G. C. Young (whose son was in the audience at the panel discussion).

The panel shied away from the simple recital of individual horror stories: the woman applicant questioned at length about provisions for birth control or child care, the prominent mathematician denied a regular position for years due to antinepotism rules, the Ph.D. relegated to typing and coffee-making chores, to concentrate on possible remedies.

A great deal of attention was focused on cultural conditioning. Young girls are indoctrinated to set low goals for themselves, e.g., to become a nurse, not a doctor, and in particular, to believe that they cannot, and indeed should not if they are to preserve their feminity, succeed in mathematics. Boys are conditioned to think of women in subservient roles also. It was pointed out that children's literature and high-school counselors are real menaces. Several feminist groups have come out with lists of literature they feel is appropriate for children, but no one has a workable proposal for dealing with the counselors. However, there are some career movies available with women in professional roles, and it was suggested that the MAA seek out women mathematicians for its films, President Victor Klee being receptive to the idea. More women lecturers in MAA, SIAM and independent projects would also contribute to the positive image and help the women students set higher goals for themselves. It is also interesting to note that men who may never have realized that there is a bias against women, and hence have through their unawareness fostered it, become cognizant of the problems and are willing to work to improve conditions for women when their daughters start thinking about career decisions.

I am the chairman of an organization which is working hard on the problem of improving the status of women in the profession — the Association for Women

in Mathematics. It is working on some of the items listed above and is trying to help women who are encountering specific difficulties due to discrimination. We are also maintaining an employment register. Moreover, last April the AMS Council established a committee on women in mathematics. I am serving on this committee, but unfortunately, as of November 1, there is still one position on the committee to be filled and its activities have amounted to the exchange of a few letters among its members.

One way to improve the status of women is to increase their visibility. For example, the panel on women was the only panel at the MAA–AMS meeting with women members. No invited speakers were women. Very few members of the MAA Board of Governors, the AMS Council and various editorial boards are women. It is not the case, as is frequently alleged, that there are no qualified women. Neither is it the case that women should be chosen for these positions because they are women, but rather because they are qualified. However, a special effort must be made to seek out qualified women because the cronyism which has operated in the past has tended to exclude women from many of these jobs, in some cases through oversight, in some through design.

Cal Moore of the University of California, Berkeley, asked from the audience whether the panel members felt that a less-qualified candidate for a job should be hired just because she is a woman. The panelists were unanimous in agreeing with Hewitt's comment that she should not be, but that her concern was that a well-qualified candidate should not be passed over just because she is a woman. In private discussions after the program, however, many women expressed the belief that at least a modified quota system is needed; that is, since past inequities have tended to reduce the pool of qualified women a certain number who are underqualified should be hired and given special consideration. (The Galbraith plan in the *New York Times*, August, 1971, describes such a system to get more women and minority group members into professional and executive positions; it includes an extensive training program.)

Many university administrators have become much more receptive to the idea of recruiting women due to pressure of investigations by the Department of Health, Education and Welfare, or the threat of such investigations. Wh.le it is not against the law to discriminate against women in faculty hiring, promotion, pay and tenure, (the Equal Pay Act excludes faculty positions), there is an executive order, number 11246 (amended by Executive Order 11375), requiring not only that those institutions and businesses holding any federal contract over 50,000 dollars not discriminate but that they have a written affirmative action plan for recruiting and promoting women and minority group members.

There was very little hostility in the questioning from the floor; in fact it was suggested that we should have planted a male chauvinist pig to liven things up. The aura of goodwill may have been due to a misplaced sense of chivalry or to lack of guts rather than to an understanding of the problems of women and a willingness

to help since several men accosted me later with such statements as "Women belong in bed, not at the board." Representatives of a more feminist point of view also surfaced later. There were those who feel that women are due reparations from such groups as the MAA and AMS because of their complicity in past inequities and those who felt that the basic reforms to improve the lot of all women should be the overriding concern rather than narrow goals of opportunities for professional achievement and recognition and for economic reward. This faction holds that time should be devoted to these societal goals rather than to proving theorems.

One mildly controversial issue did arise: what provision should be made for parttime "regular" (i.e., reasonable pay and fringe benefits and leading to tenure) appointments. Some, such as Barbara Osofsky of Rutgers, argued that the special nature of the subject makes the notion of part time employment as a mathematician impossible. Others maintained that if part time would mean the replacement of some committee work, counselling and classroom teaching by some of the traditional three K's of Kinder, Küche, and Kirche, then the concept was a useful one. Not only must women redirect their goals but men must learn to think of women as professionals. Many women feel that this process is retarded by the appeals for special considerations — parttime appointments, child care leaves, preferential class scheduling, etc. On the other hand, these same benefits should be available to men so that they may take part in the care of the family. Several questioners emphasized that nothing is more detrimental for women as a group than the existence of women doing a shoddy job. It may be argued that the bad performance of a man is not held against all men or taken as typical and therefore such inferences with respect to women are unfair. Undoubtedly, but they exist.

Women in all professions frequently assert that they are excluded from the camaraderie, the "old boy" atmosphere in which decisions are made, useful ideas are exchanged, etc. Such claims may be backed up by personal feelings, although the panelists all disclaimed such experience, but are difficult to substantiate. However, Hewitt did observe that review panels on which she has served (as the only woman) and recommendations which she has read do reflect some such atmosphere or subtle bias. Review panels certainly are part of the self-perpetuating mechanism of the male-dominated mathematical establishment. While the NSF claims to be unable to determine how many women are principal investigators on its grants, or what percentage of its reviewers for research grants are women, it does say that three percent of the entire consultant staff are women and 3.1 percent of the review panelists for fellowships are women. (These figures are for the entire agency, not just the math section.) In spite of its affirmative action plan prompted by President Nixon's directive, the agency does not seem to have acquired many women in its own upper echelons.

All panelists felt that the poor showing of women in almost all statistical analyses, e.g., the median salary of 10,000 dollars for the 2790 women mathematicians in the 1970 National Science Register vs. 15,000 dollars for the 21,610 men, cannot be

attributed solely or even primarily to discrimination. Instead, the cultural conditioning from early childhood through post-Ph.D. seems to be the chief factor operating against the potential woman mathematician. How to succeed in spite of this and how to change the conditioning to ease the path for our successors is the task to which at least some are now addressing themselves.

While the issues I have mentioned are the shared concerns of many mathematicians, men and women, the opinions and impressions are my own; they do not represent the views of the panel, the MAA or any other organization.

HISTORY IN THE MATHEMATICS CURRICULUM: ITS STATUS, QUALITY, AND FUNCTION

R. L. WILDER, University of Michigan and University of California, Santa Barbara

1. Status and quality of history. In words of that great American patriot, who contributed so magnificently to the pollution of our highways and the air we breath, viz., the late Henry Ford the First, "History is Bunk!" (Actually he said "History is more or less bunk," but like many quotations the abbreviated form is considered an improvement.) Similar sentiments were expressed by Napoleon, who characterized history as "a fraud," and Matthew Arnold, who termed it "That huge Mississippi of falsehood called history."

To judge from the present status of history in the mathematics curriculum, one might conclude that mathematicians feel much the same about history. As a rapid check, I selected 7 state institutions ranging from one of the largest universities to a small college, and 4 eminent private universities. From a search of their catalogs

---

Professor Wilder received his Ph. D. under R. L. Moore at the Univ. of Texas. He held positions at Brown Unv., Univ. of Texas, Ohio State Univ. before settling at the Univ. of Michigan for a long career up to his retirement. He is presently a Visiting Professor at the Univ. of California at Santa Barbara.

Professor Wilder has spent leaves and visits at the Institute for Advanced Study, the Univ. of Southern California, Cal. Tech., the Univ. of Colorado, U. C. L. A, and Florida State Univ.

He held a Guggenheim Fellowship, he received an Honorary Sc. D. from Bucknell Univ. and another Honorary Sc. D. from Brown University. He was the Henry Russell Lecturer, at the Univ. of Michigan for one year, and he is a member of the National Academy of Sciences. He served as President of the American Mathematical Society, 1955–56, and as President of the Mathematical Association of America, 1965–66.

Professor Wilder's main research interests are topology, foundations of mathematics, and the cultural history of mathematics.

He is the author of *Lectures in Topology* (ed. with W. L. Ayres, 1941), *Topology of Manifolds* (AMS Coll. Series, vol. 32, 1949, rev'd 1963), *Introduction to the Foundations of Mathematics* (Wiley, 1952, revised 1965), and *Evolution of Mathematical Concepts* (Wiley, 1968). *Editor.*

I determined that among the 7 state institutions, 3 have history of science departments, but no course in the history of mathematics was listed in any of these or in any of the 7 mathematics departments. Of the other 4, one listed a quarter course on the junior level ("up to the advent of calculus"), another listed a 1-semester course for teachers, another a 1-semester course in the history of elementary mathematics, and the fourth a quarter course covering material up to the 17th century and "selected topics from more recent mathematical history." Not a single one mentioned any history of modern mathematics, other than the "selected topics" cited.

After the program for this meeting was mailed out, I was pleased to receive a letter from Professor Arthur Hallerberg of Valparaiso University containing the results of a questionnaire which he had mailed out last year to 143 institutions having well-known mathematics departments. Of the 83 who replied, 41 offer no course in history of mathematics; of the rest, none requires it of mathematics majors, although 8 do require it of teaching majors. I wish I had time to include more of his results.

Further evidence may be found in the history content of the MONTHLY. By 1957, the "History" classification in the annual index had disappeared (actually it had not appeared in 1955, but was used for exactly one paper in 1956). With volume 76, 1969, under the new editorship of Harley Flanders, the index was expanded to 16 classifications. None of these is "History", which, presumably because of its rarity, appears in the classification called "General."

Now I doubt that the reason for this situation is to be found in the nature of history itself. Several reasons may be offered which seem to be more credible. In the first place, during the decline of history, mathematics itself has been undergoing rapid acceleration — some have termed it a "Golden Age." And it seems plausible that during a period of expansion in mathematical theory, interest in historical research should wane. Why bother with the past, when the future beckons so enticingly?

But I am sure this is not the whole story. To speak frankly, I have detected a current of disparagement, bordering on scorn, among research mathematicians, indicating that historical research has somewhere along the way fallen into disrepute. During the first third of the present century, it was not unheard of for a mathematics department to award a Ph.D. in history. Today the candidate for a degree in history is likely to be shunted into either the school of education, or into the department of the history of science. And department chairmen are notoriously unwilling to hire doctorates from other departments, thus adding to the unwillingness of the research-capable young man to follow up any possible interest in history.

Yet how is the student to develop an interest in history if no substantial courses are offered in it? And here is another possible clue to the decline, namely, that the courses originally taught were mostly in the history of elementary mathematics, with only brief, if any, incursions into modern history. I can't refrain from quoting from an article by the late E. T. Bell entitled "Possible projects in the history of mathematics" published in *Scripta Mathematica* in 1945 — over a quarter century ago:

"Some of the dreariest ramblings ever endured in university lecture rooms by bored students earning a fairly easy credit, were those perpetrated in the name of scientific history a generation ago by professed historians of mathematics. These well-meaning and unimaginative men transferred to history the pseudoscientific fatuity of accuracy to the sixth decimal long after a rapid succession of basic new discoveries had outmoded profitless meticulosity in science. Interest they reprobated as a vice and pedantry they lauded as a virtue, all with the supposed sanction of the scientific method, of which they were congenitally incapable of understanding anything. Their drab lectures appear to have had an unintended but predictable effect.

"Inspection of recent and current catalogues shows that the fraction of colleges, universities and teacher-training schools offering a course in the history of mathematics is negligible." (Recall that this was in 1945.)

It is clear where Bell placed the blame. And whether he was right or not I feel that, after allowing for Bell's penchant for exaggeration, he hit close to the truth.

In view of the already crowded condition of the mathematics curriculum, I know that no amount of expostulation and entreaty on my part would overcome the present apathy concerning the history of mathematics. I am convinced that only two things can enable history of mathematics to compete for a place in the curriculum. These are, first, to devise courses which will not only attract the student but be of intrinsic value to his future; and, secondly, to find a way of rejuvenating the history of mathematics so that the excitement of doing research in it will be just as great and rewarding as in mathematical research proper.

2. Function of history in the curriculum. But let me pause a moment to consider a question which I am sure some of you may be asking at this point, viz., "Why should there be more attention paid to history of mathematics? Maybe the situation is as it should be, considering how crowded the curriculum is and how difficult it is to give our students what they need for either a baccalaureate major in mathematics, or even the Ph.D. More explicitly, what function can history serve under these circumstances?"

Before I attempt to answer this, let me interpolate that if anyone had told me 30 or 40 years ago that I would one day be making a plea for history before the MAA, I would have replied "Impossible!" And my doing so is absolutely not the result of my looking around for a worthwile cause to support, or a possible title for an MAA address, but rather the attempt to answer criticisms which I have been hearing students make for years. These indicated to me there was something wrong with the present system — something which is apparently being aggravated by the recent upsurge of mathematical research. I refer to complaints from students that the various courses we are offering them are too self-centered, and that their teachers were making no effort to interrelate these courses. And they were asking, how *do* all these specialities relate to one another, and what significance do they individually have for the bulk of mathematics? Where is it all going, anyway?

A couple of years ago I participated in a CUPM panel which conducted interviews with a fairly representative group of mathematics majors, some of whom had already graduated. I was impressed by the fact that these same evidences of frustration repeatedly occurred in the criticisms made by these students.

I have often observed, too, that among some of the most capable, research-wise, of new Ph.D.'s, can often be found the greatest lack of knowledge concerning the background and significance of their work, as well as abysmal ignorance of the reasons for doing it and of the general nature of mathematics. In short, they are uneducated specialists. If you ask them why they are specialists, the best reason they can give is that this is the way to get results which merit publication and hence a good job.

Now of course they are right, and I am not one to decry specialization. In this modern day and age, we are all specialists of one sort or another. But I don't believe that courses in English literature, philosophy, or other so-called "humanities" which are commonly advocated for "broadening out" the specialist, are the answer here; their effects are too often soon smothered by the rapidly increasing burden of facts and details demanded by one's specialty. What is needed is something really germane to one's interest, which he won't forget because it really complements his interest, and which will actually be capable of serving both a humanistic purpose and a mathematical one. It should not only broaden one's outlook, showing the place of mathematics in one's culture, but it should inform him where his specialty fits into the general scheme of mathematics, how it arose in the first place, and give him a means of judging where it is likely to go. What I have in mind is the kind of knowledge about mathematics that will enable one to detect gaps where new concepts are needed; spot broad areas where new structures would provide unification and consolidation of seemingly diverse concepts; and recognize when a field has borne nearly all the mathematical fruit of which it is capable, so that it needs either to be rejuvenated by fertilization with ideas from other branches of mathematics, or possibly abandoned if its benefits to other fields are nil. The student should understand how and why the introduction of new conceptual materials may lead to the solution of long outstanding problems, as well as that once these materials are available, several working independently of one another will probably get the solution, and that he shouldn't blame himself if he was one of these and was preceded in publication. No doubt much of this kind of knowledge and perspective is acquired by experience and increasing mathematical maturity, although even in such cases I suspect that much of it is only intuitive.

3. Teaching of history. I have been pondering this situation for many years, and it is my firm conviction that the history of mathematics, when suitably conceived and adapted to the needs of our students, is precisely what is needed by many of the mathematical illiterates who pass through our departments. Now please let me make clear that I am not setting myself up as an authority on history; I am not. But the teaching of history is something with which I think all mathema-

ticians have a right to be concerned. And there are signs that this is happening. I have found out, during the past few months, that several mathematicians of excellent reputation, none of whom is a professional historian, are experimenting with history courses. One of these expressed to me the opinion that history of mathematics is "an idea whose time has come."

If this is the case, then it can be expected that there will be new workers in the field contributing their ideas regarding how it should be modernized. And I hope that my own remarks will be received in this light, viz., as a desire to contribute to the development of a history course that will perform the functions I have just mentioned.

Perhaps others share with me the feeling that we mathematicians have failed to consider the possibility of applying, to history, methods which have been so successful in the body of mathematics, viz., adopting a structural point of view. This has made it possible to consolidate and cross-fertilize seemingly unrelated parts of mathematics, thereby bringing them into a more manageable focus. Historians who must be contemplating with dismay the problem of recording all the developments of the 19th and 20th centuries might take a leaf from their mathematical colleagues' notebooks and consider whether a similar remedy might work for history.

But, you may ask, history is made by human beings, and how are you going to treat human beings by introducing higher level abstractions as we have done in mathematics? If we restrict ourselves to biographical, chronological and anecdotal details, I agree that we cannot. But if instead we treat the history of mathematics as a flow of concepts and ideas in the large, then we already raise it to the level of higher abstraction. Moreover this might make feasible the coordination and patterning of historical events in a manner quite similar to that employed in mathematics proper— but adapted to the historical point of view.

4. Cultural history. Actually, the standpoint from which I believe we should present the history of mathematics is at an even higher level than mathematics. By this I mean, to take a broad view of mathematics as a living, growing organism which is continually undergoing evolution; in short, we should study it as a culture. Only two months ago I came across a little book [1] embodying 3 lectures given in 1956 by Harry Shapiro, an anthropologist of the American Museum of Natural History. One of these lectures (the second) was devoted to the contributions which he thought the modern discovery of culture could make to historical research and writing. Although admitting that historians "have become increasingly aware of culture content," he deplored the fact that few (if any) historians "exhibit any familiarity whatever in their writing with principles that anthropologists have been able to extract from cultural data." His remarks were accompanied by examples from both Irish and American history.

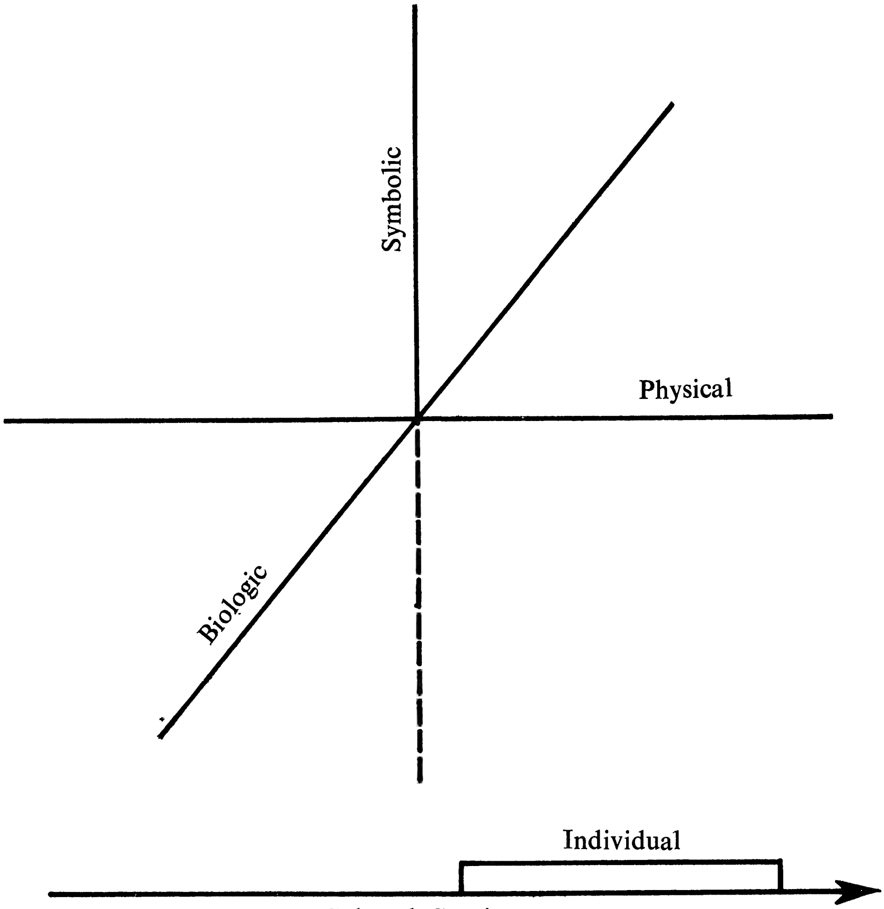
However, we cannot expect our students will have taken a course in cultural anthropology. In order to overcome this handicap, I have tried to devise a suitable substitute especially adapted to the point of view of the mathematician. This involves

making clear what is meant by a *symbol.* This is necessary since most mathematicians use the word "symbol" in a special sense, namely in the sense of so-called "mathematical symbol" or, in mathematical logic, "logical symbol." This, I have learned, has caused me to be gravely misunderstood heretofore, so I don't intend to make the same mistake now. For instance, I have been suspected of exaggerating the importance of "symboling" in view of the "glorious nonsymbolic achievements of Greek geometry" and "the Arabic development of a rhetorical algebra" [**2**]. But both Greek geometry and Arabic algebra were decidedly symbolic. My critic, a well-known historian but also a mathematician, was naturally taking it for granted that "symbol" meant "mathematical symbol" in the narrow sense.

The usual dictionary definition defines "symbol" as "something that stands for something else;" and this really sums up the matter in a nutshell. If I say the word "air" you probably think instantly of something you breathe, unless, of course, you think of one who is the beneficiary of an estate. At any rate, the word "air" stands for something else and hence is a symbol. Most words are symbols. But symbols don't have to be words; they can be traffic lights, geometric figures, finger and hand positions used by the deaf and dumb, or "peace" symbols for instance. Advertisers employ words, designs and pictures which they repeat over and over by radio, TV, print, and other forms of display, with the aim of creating symbols which will automatically pop into our minds whenever we want the sort of articles they offer for sale. "Snap, crackle, pop" is a symbol for a certain brand of cereal. It is no exaggeration to say that we are saturated by symbols. That we mathematicians customarily think of "symbols" in the narrow sense in which *we* use the term, is in itself an indication of how specialized we have become in our thinking.

Once we have learned what a symbol stands for, we usually develop a "habit" attitude toward it. An experienced driver habitually stops his car when he comes to a red light or a "STOP" sign; it isn't necessary for him to pause to inquire the meanings of these symbols. Indeed, for many symbols we get into the habit of treating them as though they were identical with their meanings — which leads to great efficiency but can be dangerous sometimes. In such a context they function only as *signs.* Animals other than man understand and react to signs. But they cannot, apparently, create symbols. To create a symbol, or as I shall say, *to symbol* (see [**3**]), one must be able to assign to some combination of sounds, events, structure, or other thing capable of being perceived, a *meaning.* We can teach a dog to follow closely at our heels on the command "heel!" But it is we, as humans, who invented this signal; the dog did not invent it and to him it is only a sign to be reacted to in a fashion to which he has been trained. Similar remarks can be made about chimps who professedly "count" up to 7. The experimenter assigns the meanings to the lights or colors which serve as the symbols, not the chimp. To use a biological term, the ability to symbol is *species-specific* (see [**4**], for instance), and can be used to distinguish humans from other animals; it is a necessary and sufficient condition for being a member of the species *homo sapiens.*

Diagram No 1



To teach the mathematics student the meaning of the word "culture," we can now proceed as follows: Consider diagram number 1. Aping Euclid, this is supposed to represent the world in which we live—only now it is the world of culture. Euclid's 3-dimensional world has been compressed into the one axis, labelled "Physical." Everything, living or not, has a physical form, but if a living thing, it has a place in the biological realm and is not confined to the 1-dimensional in this scheme, but has another degree of freedom in the plane of biological forms. But when we, as human beings, use our faculty of symboling in order to conceptualize, we then are enabled to enter a new dimension, not accessible to other life forms; this is the world of culture. Without symbols we could not enter it. The world in which we live is compounded of tools and technology, rituals and beliefs, architecture and the arts, literature and the sciences — including, of course, mathematics. All of these are based on symbols,

without which we would have no words to communicate, or with which to hand on to our progeny the vast conceptual world that we have created — a world which molds our beliefs, customs, and language as we are reared in our particular niche of this world of culture (see Note 1).

This reminds us that the world of culture is not a static thing; it is continually undergoing expansion and change. So another dimension should be added to the diagram. For sake of simplicity, I have separated this — having now compressed the world of culture into a single dimension — and by a single line represented the flow of culture, or the "cultural continuum" as it is frequently called by the anthropologist. The individual is introduced into this flow when he is born, is culturally conditioned while young, and eventually contributes to the cultural environment through his own inventions and creations, and ultimately dies. His ability to make his contributions was conditioned by his physical, biologic (or genetic) and cultural heritage; if he is poorly endowed with any of these, his contribution may be little or nil. Genetically he may be a genius, but if he is born into a culturally poor area of the world of culture, that genius may never show. But whatever he accomplishes will be dependent upon the labors of those who preceded him, and which reach him via the written or spoken word, i.e., by symbols.

But I must skip details. We are familiar with the fact that various forms evolved in the physical world, and later, living forms evolved. But the evolutionary process did not stop there. Just as the evolution of the living cell made possible the complexity of life forms familiar to the biologist, so did evolution of the ability to symbol in the species *homo sapiens* make possible the complexity of cultures that we see today. And just as the history of living forms could be expanded and made more meaningful by the Darwinian and post-Darwinian theory of evolution, so can the cultural history of man — and this includes the history of science and its subdomain, the history of mathematics — be supplemented by a theory of evolution. As was made clearly evident at the centenary celebration, in 1959, of the publication of Darwin's *Origin of Species* (whose proceedings have been published in 3 volumes [5]), modern anthropology has come to recognize that the evolutionary process did not stop with the biological, but continued with the cultural. The evolution of culture has become quite as active a field of investigation as has been the evolution of biological forms. And, I might add parenthetically, cultural evolution has been accompanied by virtually the same sort of disagreement in various scholarly circles as was the theory of biological evolution. This is why we still use the term "theory" in connection with it, although since it explains so many things that otherwise appear to have only vague mythical or philosophical explanations, there seems little doubt of its scientific utility and respectability.

I would like to suggest that a semester course in what I call "Evolution of Mathematical Concepts and Theories" will provide the student with answers to such questions as "How did mathematics get this way?" and inform him of what he is

likely to see in the future. Such a course would be based on history, but history in the sense of a continuously evolving subculture. The history involved could be either ancient or modern, or both, depending on the mathematical maturity of the students. It need not replace the more orthodox type of history course for the history major, although even he should profit by taking it before his other courses in history.

5. History as evolution. Now the history would provide principally the *stages* of the evolutionary process. But there is more to evolution than these. If we take a look at what the biologist has done, we shall notice that some of the major problems of biological evolution have been concerned with the dynamics of the process; i.e., with those forces that were instrumental in producing the stages. Darwin himself proposed the theory of natural selection — a survival of the fittest. Later biologists discovered gene shuffling and mutational forces. But probably due to its late arrival on the scientific scene, the theory of cultural evolution seems not to have advanced so far (see Note 2). Anthropologists have been unable to agree on the stages of general cultural evolution — we must recall that they must rely heavily (in addition to data on existing primitive cultures) on archaeological rather than on recorded evidence, and culture is simply not found in diggings but must be inferred from pots, bones, weapons and other physical evidence. Certain forces have, to be sure, been discovered, such as *diffusion* — the passing of such cultural elements as customs, religions and tools from one culture to another. But even here, much time and energy has been consumed in arguments over whether diffusion or independent invention accounted for similarities between different cultures. It has come to be recognized, however, that these are not mutually exclusive. For instance, counting probably originated independently in many different cultures, but once a primitive tribe comes in contact with a more advanced stage of civilization, diffusion of the counting practices from the more advanced to the less advanced usually occurs.

In the "Points for Discussion" for a panel on *Social and Cultural Evolution* during the Chicago Darwinian Centenary which I mentioned a while ago, can be found the following [**5**; vol. 3, p. 233]: "As to the macrodynamics of cultural evolution, its causes and principles, ... there is as yet no general agreement. For the near future this subject needs careful research. This is necessary as a basis for any attempt to predict or control the direction of cultural evolution."

Fortunately in mathematical history we have a wealth of recorded information. I use the word "wealth" in spite of the fact that historians bemoan the loss of most Greek mathematical works, for example. In comparison with the scarcity of early remains which the anthropologist has to work from, we are indeed lucky. It would be nice to know more about how counting and the number concept evolved, and just what individuals were responsible for the geometric discoveries and inventions presented to us in finished form in Euclid's *Elements*. But we should be grateful that we can infer pretty well just what the general outline of early mathematical development was like and of course in the case of modern mathematics, we indeed

have a wealth of recorded material. Regarding the stages through which mathematics has passed there is still some conjecture, especially on the elementary level. As to the forces involved, there seems little reason to think that they were much different (except for being fewer in number) from forces operating today.

In diagram number 2, early *stages* in the evolution of number are listed (see **[6]**, p. 180). I must omit details. The first two stages we get from the anthropologist. Comparison by (1-1)-correspondence can be inferred from anthropological evidence regarding early number words, and tallying is evidenced in many early numerical records — the earliest being the find in 1937 of the radius of a young wolf from paleolithic times which is covered with notches so grouped as to be indubitably a tally

Diagram No. 2

Stages in Evolution of Number

One-two differentiation	Numeral Systems
One-two-many	Mysticism
Comparison: (1-1)-Correspondence	Operations with numerals
Tallying	Fractions
Number words	Zero
Ideographs	Negative, complex numbers
	Etc.

(see Note 3). The only item about which there may be some question is "Mysticism" (Note 4). Certainly most of us are familiar with Pythagorean numerology, but it had its counterpart in early Babylonian mysticism and I believe there is good reason for assigning it a part in the evolution of the concept of number — in short, with numbers becoming nouns, or things. It survives today, of course, in the host of numerologists, astrologers, and number lore. The number 13 has such a bad reputation as to induce many modern hotels to omit the 13th floor, although I am sure that their managers could not tell what the number 13 is as a concept. In fact, all of these stages have their modern counterparts, just as many early biological forms still exist in modern form.

Diagram No. 3

Forces of Mathematical Evolution

1. Environmental Stress	6. Generalization
(a) Physical	7. Consolidation
(b) Cultural	8. Diversification
2. Hereditary Stress	9. Specialization
3. Symbolization	10. Cultural Lag
4. Diffusion	11. Cultural Resistance
5. Abstraction	12. Selection

6. Forces of evolution; how mathematics grows. By way of contrast, consider the list of *forces* of mathematical evolution given in diagram number 3 (Note 5). Again I shall omit details, but shall briefly illustrate their nature. (See, however, the discussion in my book referred to above.)

Environmental stress is listed first, since it was unquestionably the first and most elementary of the forces involved in the evolution of mathematics. Indeed, it was likely active even before man evolved, since capability of one-two differentiation can be exercised by most animals and is not necessarily cultural in nature. In order to adapt, the animal must be able to sense whether he is facing one or more enemies, for example. Thus much of the initial environmental stress was physical in nature. However, with the evolution of culture in man, environmental stress of a cultural nature began to play a part, as might be expected since man was entering a new world. Comparison by matching, tallying, and, eventually, the invention of number words took place. And, when urban life evolved, the stress exerted by building, architecture, imposition of taxes and recording thereof and the like forced the invention of elementary calculating. And of course cultural stress still plays an active part in mathematical evolution, as those who were affected by the demands of the second world war can testify. And don't think that present economic conditions resulting in a lack of jobs for new Ph.D.'s is not going to have its effect!

I shall comment only briefly on *how* these forces individually work – actually I have not had time in my own studies to complete such an analysis (no geneticist has solved all the problems concerned with mutation). But I am sure that even superficial consideration of them will be enough to indicate their general function and importance. Symboling was already active in the invention of number words; as the mathematician has often done, primitive man first utilized words of ordinary discourse, as in the use of "hand" for the number 5 for instance. L. L. Conant's classic work of 1896, "The Number Concept," is revealing here [7]. And of course symboling is one of our chief tools, as are also abstraction and generalization. Diffusion, cultural lag and cultural resistance I have borrowed from the anthropologist. Diffusion I have already defined earlier; we wouldn't be using the Babylonian sexagesimal system for fractional measurement of angles if it hadn't diffused from one ancient culture to another, and, eventually, into our own Western culture. Even our journals can be considered as a means of diffusion of mathematical ideas.

Cultural lag can be thought of as a sort of "laziness," or indisposition to make the effort to adopt a more efficient tool. I just mentioned our use of Babylonian numeration in angle measurement, and I imagine cultural lag also played some part in this, although I'll leave that to the professional historian. A current example may be found in the plans for converting to the metric system in this country; the big problem will be overcoming cultural lag. Cultural resistance is a more overt obstacle to diffusion. Most missionaries have encountered it, and for a whole century the English mathematical community resisted adopting the Leibnizian differential notation presumably out of loyalty to Newton. I am sure some of you can recall

instances of cultural resistance in mathematical circles, in cases where one group of mathematicians refuses to adopt more efficient methods and concepts which have evolved in other groups; of course cultural lag may be operative in such instances also.

Two of the most important and profound of the forces listed are hereditary stress and consolidation. Only as mathematics has become more mature and complex has their influence become so great as to render them obvious. Hereditary stress is a cultural stress created by the accumulation, usually over a period of extended duration, of concepts and their interactions *within* a system. I find that historians have sometimes detected it. For example, the late historian of science, George Sarton [**8**; p. 444], stated: "The whole fabric of science seems . . . to be growing like a tree; in both cases the dependence upon the environment is obvious enough, yet the main cause of growth — the growth pressure, the urge to grow — is *inside* the tree, not outside [italics ours]." I believe, too, that what Struik has suggested [**9**] as a cultural force and called "cultural impetus," is largely a part of hereditary stress (although sometimes cultural stress of environmental type). Hereditary stress was active in the ultimate admission of complex numbers to mathematical respectability, although for a long time they were what Cardan termed *numeri ficti*, or *numeri falsi*. A prime example in modern mathematics is set theory which was born from the demands of the theory of functions. As each of us is introduced by his mentors into the mathematical culture stream, we inevitably react to hereditary stresses by recognizing where improvements, new theorems, and new concepts will contribute to the growth of the branch of mathematics in which we have elected to work. The psychological aspects of our reactions have been described by both Poincaré and Hadamard.

Although it is one of the most active forces in mathematics today, consolidation has operated throughout mathematical history. As far back as old Babylon, when the Akkadians conquered Sumer, they consolidated the old Sumerian terms for "multiplies by," "find the reciprocal of," with their arithmetic in the form of ideograms, thus initiating an important advance in mathematical symbolism. Derek Price cites the consolidation in Ptolemy's *Almagest* of the Greek *geometric* astronomy with the Babylonian *numerical* astronomy as the probable reason why Western science has reached such heights while this did not occur in other civilizations, such as China, which had the ingredients for such an achievement. He makes out quite a convincing case for this thesis in the first chapter of his book "Science since Babylon" [**10**].

Coming nearer to the modern era, an outstanding example of consolidation was that of number with line, as a result of which the analysts preceding the so-called "Arithmetization of analysis" were able to create a large body of good mathematics with the help of geometric intuition. And during the modern era, one of the most interesting examples was that of the consolidation of algebra and topology. Such fields as algebraic geometry, differential geometry, differential topology were formed by consolidation. It can be inferred, that as the body of mathematics grows, opportu-

nity for consolidation increases, and the greater power that is thus achieved can be seen in the solution of problems which had defied solution in their own fields. The process effects a kind of cross-fertilization.

It should be noticed that generally these forces do not act independently. Much as in biology, where adaptation often joins with gene mutation to effect survival, so in mathematics consolidation is frequently forced by hereditary stress; and in the process, diffusion, generalization and abstraction may play a part. It was the consolidation of the group-theoretic features of various mathematical theories that led to abstract group theory, and category theory is a nice example where generalizing from the features of the plethora of homology theories of modern algebraic topology resulted in a consolidation of common elements which is proving one of the most important modern tools in modern mathematics. If this sort of thing did not happen mathematics would simply grow like a tree with innumerable branches having no contacts with one another, with eventual chaos as the probable outcome.

7. Example of a course. It is impossible for me to make, in 50 minutes, the complete case for what I firmly believe is an area that offers much promise for research. I shall conclude with some comments on what I think can be done for the student on the basis of these ideas. First let me briefly exhibit some outlines indicating the nature of a one-quarter course I gave at the University of California in Santa Barbara a year ago. Diagram number 4 gives a list of the general topics covered. The students were supposed to be juniors and seniors, but a number of graduates were allowed to attend, including one who was working on the Ph.D. in philosophy. Since some of these topics may seem strange, I will exhibit outlines for two of them.

Diagram No. 4

A Course Outline

1. Symbols and symboling
2. Culture
3. Counting
4. Evolution of counting
5. Evolution of geometry
6. Evolution of real number system
7. Aspects of reality
8. Evolution of function, set concepts
9. Evolution of real analysis
10. Emergence of contradictions
11. Identification, analysis of evolutionary forces
12. Role of individual in evolution
13. Philosophies of mathematics
14. Evolutionary "laws"

Diagram number 5, "Aspects of Reality," may be roughly explained by pointing out that throughout the course, I repeatedly emphasized, as opportunity offered, that as a part of the world of culture, mathematics is just as real as any part of the physical world. But since it has a tendency to deal in ever higher levels of abstraction, we continually need reassurance that our creations do really add to the existing body of

mathematical reality. This has led to the use of models — which will explain why several of the items relate to model theory.

Diagram No. 5

7. Aspects of Reality

(a) Physical; perception of
(b) Extension to cultural environment
(c) Inception of use of models
  (i) Function to maintain contact with reality
  (ii) Beltrami, Klein models
(d) Evolution of model theory
(e) Role of models in axiomatics
(f) Mathematical reality
  (i) Reality of concepts after adoption by mathematical community

Referring to diagram number 6: The evolution of function and set was chosen as one of the topics partly because I could count on everyone having some acquaintance with these notions, and partly because they also offered an excellent example to show the interplay of the evolutionary forces.

Diagram No. 6

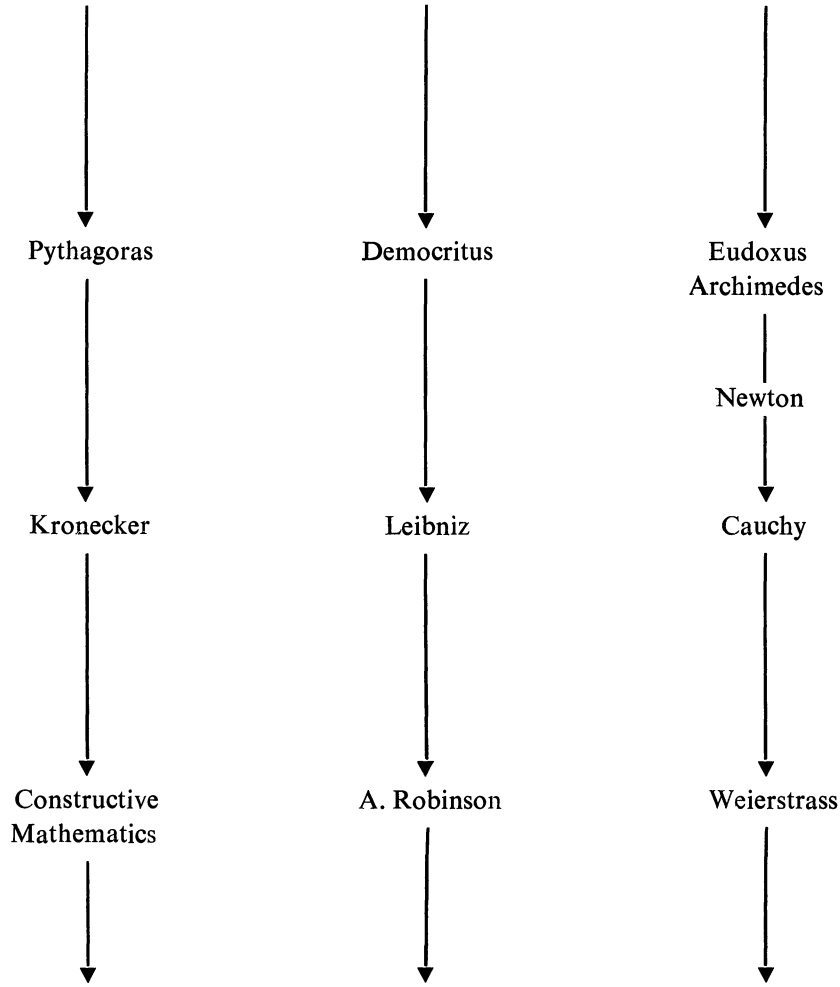
8. Evolution of function and set concepts

(a) Theory of sound; vibrating string
(b) D'Alembert: Euler, Bernoulli solutions
  (i) Disagreement over meaning of "function"
(c) Theory of heat; Fourier
  (i) "Uninhibited" notion of function
(d) Dirichlet's conditions
(e) Riemann's work on trigonometric series
  (i) Integrability conditions
  (ii) Influence on function concept
(f) Cantor's uniqueness theorem
  (i) Species of a point set
  (ii) Inception of set theory
(g) Emergence of new principles
  (i) Continuum hypothesis; axiom of choice

In showing how the processes of evolution work, I made extensive use of charts or diagrams, to show graphically the flow of influences of one part of mathematics upon another, as well as consolidations. Most of these are too complicated to squeeze into a compact diagram. Here is one (diagram number 7) containing some elements of conjecture — indulgence in reflecting on "What might have happened" was not frowned on, by the way. I chose to exhibit this one today because it is so simple (not historically complete, but purely indicative) and is somewhat topical in view of the subject of the lectures which Professor Robinson is giving at this meeting. Professor Robinson has discussed in his book [11] some of the reasons why the path depicted by the middle column was not pursued by analysis; the right-hand column represents the actual course of analysis, it will be observed.

I like to think that those who took the course acquired some understanding of how the various courses they were taking came into being, and how they were interre-

Some possible directions for the foundations of Analysis



lated — although I left much of this to the individual to reason out for himself using the ideas he had, hopefully, assimilated. Certainly each understood that mathematics is still undergoing evolution, and that if he was going to make it a career, his only chance for success was to enter the stream at some likely point of his own choice; but to expect that he would have to spend much of his future in keeping up with the changes that would inevitably occur.

Obviously this was not an orthodox history course. It was more in the nature of what the historian of science would call a science of the history of mathematics. If may be that a history course along more orthodox lines can be devised which will accomplish much the same ends in a more efficient manner. I have been pleased, during the course of preparing this material, to hear from several mathematical col-

leagues who are working on the problem of a suitable modern history course — so much so, that I earnestly look forward to the rejuvenation of history in a more up-to-date form in the classroom; and even that the subject will reach such a degree of acceptance as to be again considered worthy of the Ph.D. in mathematics.

8. Philosophical implications. One final word: When I was briefly discussing mathematical reality, perhaps some of you wondered where Platonism fits in? In particular, does a theory of mathematical evolution, based on the location of mathematical reality in the world of culture run counter to Platonism? The answer is emphatically "No"; no more than Darwinism destroyed existing religions, despite the fears of the clergy. The anthropologist studies religions as a part of culture; to him they form an adapting mechanism, and he takes no position, as a scientist, on whether they represent a reality outside the world of culture or not. Similarly, a theory of mathematical evolution can study, using the tools of science, the manner in which Intuitionism, Formalism, Constructivism, Platonism, or any other philosophy of mathematics evolved. But it takes no position on their so-called "Truth," or on what other possible types of reality they may represent. So if you are a Platonist, go ahead and enjoy it!

Except for minor changes and addition of literary references, this is a verbatim copy of the author's address before the summer meeting of the Association at Pennsylvania State University, 1971.

NOTES

1. See Ernst Cassirer, An Essay on Man, Yale Univ. Pr., New Haven, Conn., 1944. "As compared with other animals man . . . lives . . . in a new *dimension* of reality. . . . Physical reality seems to recede in proportion as man's symbolic activity advances. . . . He has so enveloped himself in linguistic forms, in artistic images, in mythical symbols or religious rites that he cannot see or know anything except by the interposition of this artificial medium" (ibid., p. 25).

2. See, however, L. A. White, Energy and the evolution of culture, Amer. Anthropologist, vol. 45 (1943), pp. 335–356, for a proposal regarding general cultural evolution and the forces governing it; also W. F. Ogburn, On Culture and Social Change, O. D. Duncan, ed., Univ. of Chicago Pr., 1964.

3. See the note in Isis, vol. 28 (1938), pp. 462–463, referring to a news item in the Illustrated London News of Oct. 2, 1937, concerning excavations made by Karl Absolon in Czechoslovakia.

4. Whether passage through a stage in which different numeral forms were used for various categories of objects and concepts, is conjectural, although there is much evidence for it. For instance, this phenomenon occurred among certain Plains Indian tribes, as well as among Northwest Indian tribes and other cultures; remains of such a classificatory numeral system are found in the Japanese language.

5. Except for the addition of "Specialization," this is the list of forces given on p. 169 of my book [6].

Bibliography

1. Harry L. Shapiro, Aspects of Culture, Rutgers Univ. Pr., 1956.

2. Carl B. Boyer, The anthropology of mathematics, Science, vol. 163 (1969) p. 799.

3. L. A. White, Symboling: A kind of behavior, The Jour. of Psychology, vol. 53 (1962) pp. 311–317.

4. E. H. Lenneberg, On learning language, Science, vol. 164 (1969) pp. 635–643.
5. Sol Tax ed., Evolution after Darwin, 3 vols., Univ. of Chicago Pr., Chicago, Ill., 1960.
6. R. L. Wilder, Evolution of Mathematical Concepts, John Wiley and Sons, N.Y., 1968.
7. L. L. Conant, The Number Concept, Macmillan, N.Y., 1896.
8. G. Sarton, Science and morality, in Moral Principles of Action, ed. Ruth N. Anshen, Harper and Row, N.Y., 1952.
9. A review of [6], this Monthly, vol. 76 (1969), pp. 428–429.
10. Derek J. de Solla Price, Science since Babylon, Yale Univ. Pr., New Haven, Conn. 1961.
11. A. Robinson, Non-standard Analysis, North-Holland Pub. Co., Amsterdam, 1966.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306; notes are usually limited to three printed pages.

VARIATIONS ON THE BINOMIAL SERIES

H. POLLARD, Harvard University and Purdue University and

O. SHISHA, Aerospace Research Laboratories, Wright-Patterson AFB, Ohio

1. Introduction. This study began when one of us asked the other whether there exists a reasonable continuous analog of the equality

$$(1) \qquad (1+z)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k \quad (\alpha > -1, |z| = 1, z \neq -1),$$

where, for real u,

$$(2) \qquad \binom{\alpha}{u} = \frac{\Gamma(\alpha+1)}{\Gamma(u+1)\Gamma(\alpha-u+1)}.$$

It is natural to try to replace the right hand side of (1) by $\int_0^\infty \binom{\alpha}{u} z^u \, du$. This, however, led us up a blind alley.

We then observed that from a known formula (§3) one obtains

$$(3) \qquad \binom{\alpha}{u} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iut}(1+e^{it})^\alpha dt, \ \alpha > -1, \ -\infty < u < \infty.$$

In particular, for k an integer, $\binom{\alpha}{k}$ is the kth Fourier coefficient of $(1+e^{it})^\alpha$. Since by (2), $\binom{\alpha}{k} = 0$ for $\alpha > -1$ and $k = -1, -2, \cdots$, we can interpret (1) as an equality (throughout $(-\pi, \pi)$) between $(1+e^{it})^\alpha$ and its Fourier series $\sum_{k=-\infty}^{\infty} \binom{\alpha}{k} e^{ikt}$. Therefore, a continuous analog of (1) appears to be obtained by inversion of the Fourier transform (3):

$$(1+e^{it})^\alpha = \int_{-\infty}^{\infty} \binom{\alpha}{u} e^{iut} du, \quad -\pi < t < \pi,$$

namely,

$$(1+z)^\alpha = \int_{-\infty}^{\infty} \binom{\alpha}{u} z^u du \qquad (\alpha > -1, |z| = 1, z \neq -1). \tag{4}$$

Relation (4) (to which we return in §3) could hardly have been anticipated from (1). It was given (in a less compact notation) by S. Ramanujan [**1, 2**].

2. A simple observation. *Let a (complex) function f belong to $L[-\pi,\pi]$, let $-\pi \leqq x-\delta < x+\delta \leqq \pi$, and suppose that f is of bounded variation in $[x-\delta, x+\delta]$. Then*

$$\lim_{R\to\infty} \frac{1}{2\pi} \int_{-R}^{R} e^{-iux} a(u) du = \frac{f(x^+)+f(x^-)}{2} = \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=-N}^{N} e^{-nx} a(n), \tag{5}$$

where

$$a(t) = \int_{-\pi}^{\pi} e^{iut} f(u)\, du, \quad -\infty < t < \infty. \tag{6}$$

(5) follows at once from Jordan's tests for Fourier series and Fourier transforms, and shows that the expansion

$$\frac{f(x^+)+f(x^-)}{2} = \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=-N}^{N} e^{-inx} a(n) \tag{7}$$

has the continuous analog:

$$\frac{f(x^+)+f(x^-)}{2} = \lim_{R\to\infty} \frac{1}{2\pi} \int_{-R}^{R} e^{-iux} a(u)\, du.$$

3. Variations on the binomial series. As an example of (5), consider the function $f(t) \equiv (1+e^{it})^\alpha$, $-\pi < t < \pi$, where α is a constant > -1 (we always take for a power its principal value). It is not difficult to see that $f \in L[-\pi,\pi]$. Let $-\pi < x < \pi$. Then by (5),

$$(1+e^{ix})^\alpha = \lim_{R\to\infty} \frac{1}{2\pi} \int_{-R}^{R} e^{-iux} a(u)\, du = \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=-N}^{N} e^{-inx} a(n), \tag{8}$$

where for $-\infty < t < \infty$,

$$\begin{aligned} a(t) &= \int_{-\pi}^{\pi} e^{iut}(1+e^{iu})^\alpha du = \int_{-\pi}^{\pi} e^{iut}\left[2e^{iu/2}\cos\left(\frac{u}{2}\right)\right]^\alpha du \\ &= 2^\alpha \int_{-\pi}^{\pi} e^{iu[t+(\alpha/2)]} \cos^\alpha\left(\frac{u}{2}\right) du = 2^{\alpha+1} \int_{-\pi/2}^{\pi/2} e^{i\tau(2t+\alpha)} \cos^\alpha \tau\, d\tau \end{aligned}$$

$$= \frac{2\pi\Gamma(\alpha+1)}{\Gamma(\alpha+t+1)\Gamma(1-t)},$$

the last equality following from a known formula [3, (7.6.1)]. Thus by (2),

$$a(t) = 2\pi \binom{\alpha}{-t}, \quad -\infty < t < \infty. \tag{9}$$

Consequently, from (8),

$$(1+z)^\alpha = \lim_{R\to\infty} \int_{-R}^{R} \binom{\alpha}{u} z^u du = \lim_{N\to\infty} \sum_{n=-N}^{N} \binom{\alpha}{n} z^n, \quad |z| = 1, z \neq -1. \tag{10}$$

Since $\binom{\alpha}{k} = 0$ for $k = -1, -2, \cdots$, we have from (10) the binomial relation

$$(1+z)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n \qquad (|z| = 1,\ z \neq -1).$$

The first equality in (10) gives the continuous analog

$$(1+z)^\alpha = \lim_{R\to\infty} \int_{-R}^{R} \binom{\alpha}{u} z^u du \qquad (|z| = 1,\ z \neq -1).$$

Actually, the last limit can be written, for the z's in question, as an improper Riemann integral $\int_{-\infty}^{\infty} \binom{\alpha}{u} z^u du$, which converges absolutely if $\alpha > 0$ but not if $-1 < \alpha \leqq 0$ [3, §7.6].

Another modification of the binomial series is obtained by considering the function $f(t) \equiv (1+e^{it})^\alpha e^{-ict}$, $-\pi < t < \pi$, where $\alpha(> -1)$ and c are real constants. Let $-\pi < x < \pi$; then by (5) and (9),

$$f(x) = \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=-N}^{N} e^{-inx} a(n),$$

where, for every real t,

$$a(t) = \int_{-\pi}^{\pi} e^{iu(t-c)}(1+e^{iu})^\alpha du = 2\pi \binom{\alpha}{c-t}.$$

Thus

$$f(x) = \lim_{N\to\infty} \sum_{n=-N}^{N} \binom{\alpha}{n+c} e^{inx},$$

and we have the following result: *Let* $\alpha(> -1)$ *and* c *be real constants. Then*

$$(1+z)^\alpha = \lim_{N\to\infty} \sum_{n=-N}^{N} \binom{\alpha}{n+c} z^{n+c}, \quad |z| = 1,\ z \neq -1. \tag{11}$$

Observe that, by (4),

$$(1+z)^\alpha = \int_{-\infty}^{\infty} \binom{\alpha}{u+c} z^{u+c} \quad (\alpha > -1, -\infty < c < \infty,\ |z| = 1,\ z \neq -1),$$

a continuous analog of (11) in which the infinite sum is replaced by an improper Riemann integral.

4. Another example. Taking, in (5), $f(x) \equiv 1$ does not yield a pair of relations, one of the pair being a continuous analog of the other. For, while the first equality in (5) yields for this f,

$$\frac{2}{\pi}\int_0^\infty \cos(xu)\,\frac{\sin(\pi u)}{u}du = 1, \quad -\pi < x < \pi, \tag{12}$$

the second equality there reduces to $1 = 1$.

Taking, in (12), $x = 0$ and making a simple substitution yield the familiar formula

$$\operatorname{sgn} y = \frac{2}{\pi}\int_0^\infty \frac{\sin(yt)}{t}dt, \quad -\infty < y < \infty, \tag{13}$$

where $\operatorname{sgn} y$ is the "sign of y", namely, 1 if y is positive, -1 if $y < 0$, and 0 if $y = 0$.

To arrive at a discrete analog of (13), take, in (5), $f(x) \equiv \operatorname{sgn} x$. Then the second equality in (5) readily gives the well-known relation

$$\operatorname{sgn} x = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)x]}{2n-1}, \quad -\pi < x < \pi.$$

If $a > 0$, then for $-\pi/a < x < \pi/a$,

$$\operatorname{sgn} x = \operatorname{sgn}(ax) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)ax]}{2n-1}.$$

In particular, taking $a = \frac{1}{2}$ yields

$$\operatorname{sgn} x = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(n-\frac{1}{2})x]}{n-\frac{1}{2}}, \quad -2\pi < x < 2\pi.$$

The first equality in (5), for $f(x) \equiv \operatorname{sgn} x$, does not lead to any further representation of this function.

5. The function $a(t)$ and the sequence $a(n)$. In the example

$$f(t) \equiv (1 + e^{it})^\alpha, \quad \alpha > -1,$$

the $a(n)$ in (5) is $2\pi\binom{\alpha}{-n}$, and the $a(t)$ there is the natural interpolatory function to the sequence $(a(n))_{n=-\infty}^{\infty}$, namely, $2\pi\binom{\alpha}{-t}$. Such a simple situation does not always occur. Consider, for example, the exponential series $\sum_{k=0}^{\infty} z^k/k!$, converging (uniformly) to e^z on the unit circumference $|z| = 1$. If $-\pi < x < \pi$, then

$$\lim_{N\to\infty}\frac{1}{2\pi}\sum_{n=-N}^{N} e^{-inx}\frac{2\pi}{\Gamma(n+1)} = e^{e^{-ix}},$$

and

$$\frac{2\pi}{\Gamma(n+1)} = \int_{-\pi}^{\pi} e^{iun}e^{e^{-iu}}du, \quad \text{for } n = 0, \pm 1, \pm 2, \cdots.$$

Thus, for $f(t) \equiv e^{e^{-it}}$, the $a(n)$ in (5) is $2\pi/\Gamma(n+1)$. But for this f, the function $a(t)$ is not $2\pi/\Gamma(t+1)$, for the equality

$$\lim_{R\to\infty} \frac{1}{2\pi} \int_{-R}^{R} e^{-iux} \frac{2\pi}{\Gamma(u+1)} du = f(x)$$

fails for $x = 0$, for which the last limit is ∞. Indeed, whereas $|a(t)|$ is bounded and converges to 0 as $t \to -\infty$,

$$\left|\frac{2\pi}{\Gamma(t+1)}\right| \equiv 2\left|\Gamma(-t)\sin(\pi t)\right|$$

takes on, for negative t, arbitrarily large values.

One can, however, for every $f \in L[-\pi,\pi]$, relate the function $a(t)$ to the sequence $(a(n))_{n=-\infty}^{\infty}$. In fact, the following holds:

Let f be a complex function belonging to $L[-\pi,\pi]$ and set

$$a(t) = \int_{-\pi}^{\pi} e^{iut} f(u)\, du, \quad -\infty < t < \infty.$$

Then, for every real t,

$$a(t) = \lim_{N\to\infty} \sum_{n=-N}^{N} \frac{\sin[\pi(t-n)]}{\pi(t-n)} a(n),$$

where, for $n = t$, the last ratio is to be understood as 1.

Proof. We may assume t is not an integer. The Fourier series of f may be multiplied by the function e^{ixt}, which is of bounded variation in $[-\pi,\pi]$, and integrated term by term to yield $\int_{-\pi}^{\pi} f(x)e^{ixt}dx$. Namely,

$$\begin{aligned} a(t) &= \int_{-\pi}^{\pi} f(x)e^{ixt}dx = \lim_{N\to\infty} \sum_{n=-N}^{N} \int_{-\pi}^{\pi} \frac{1}{2\pi} a(-n)e^{inx}e^{ixt}dx \\ &= \lim_{N\to\infty} \sum_{n=-N}^{N} \frac{a(n)}{2\pi} \int_{-\pi}^{\pi} e^{ix(t-n)}dx = \lim_{N\to\infty} \sum_{n=-N}^{N} \frac{\sin[\pi(t-n)]}{\pi(t-n)} a(n). \end{aligned}$$

The authors wish to thank the referee and Professors Y. Katznelson and D. J. Newman for their valuable suggestions.

References

1. S. Ramanujan, Some definite integrals, Proc. London Math. Soc. (2), 17 (1918), Records for January 17, 1918.

2. ———, A class of definite integrals, Quarterly J. Math., 48 (1920) 294–310.

3. E. C. Titchmarsh, Introduction to the Theory of Fourier Integrals, 2nd Edition, Oxford University Press, 1948.

ON THE GREATEST ORDER OF AN ELEMENT OF THE SYMMETRIC GROUP

M. B. NATHANSON, University of Rochester

Let $f(n)$ denote the greatest order of a permutation in the symmetric group S_n. Since S_n is isomorphic to a subgroup of S_{n+1}, it follows that $f(n) \leqq f(n+1)$, and so f increases monotonically. Using the prime number theorem, Landau [2, p. 225] showed that $\log f(n)$ is asymptotic to $\sqrt{n \log n}$, and Shah [3] has slightly strengthened this result. In this note I give an elementary proof that $f(n)$ grows faster than any power of n.

All lower case latin letters stand for integers. The greatest common divisor and least common multiple of $a_1, \cdots, a_k$ are denoted $(a_1, \cdots, a_k)$ and $[a_1, \cdots, a_k]$ respectively.

LEMMA 1. *For all positive integers n,*

$$f(n) = \max \{[a_1, \cdots, a_k] \mid n = a_1 + \cdots + a_k$$

and $a_i > 0$ for $i = 1, \cdots, k\}$.

Proof. It is well known [1, Theorems 5.1.1 and 5.1.2] that any permutation σ in S_n can be written as the product of disjoint cycles of lengths $a_1, \cdots, a_k$, where $n = a_1 + \cdots + a_k$, and that the order of σ is $[a_1, \cdots, a_k]$. Conversely, for any partition of n as the sum of positive integers, $n = a_1 + \cdots + a_k$, there is a permutation σ in S_n which is the product of k disjoint cycles of lengths $a_1, \cdots, a_k$.

LEMMA 2. *For any positive integers $a_1, \cdots, a_k$,*

$$\prod_{i=1}^{k} a_i \leqq [a_1, \cdots, a_k] \prod_{1 \leqq i < j \leqq k} (a_i, a_j). \tag{1}$$

Proof. Let p be a prime number, and let s_i be the exact power of p that divides a_i. Clearly, we can arrange the a_i so that $s_1 \leqq s_2 \leqq \cdots \leqq s_k$. The exact power of p dividing $[a_1, \cdots, a_k]$ is s_k and the exact power of p dividing $\prod_{1 \leqq i < j \leqq k} (a_i, a_j)$ is $\sum_{i=1}^{k-1} s_i(k-i)$. Therefore, the power of p dividing the right-hand side of inequality (1) is

$$\sum_{i=1}^{k-1} s_i(k-i) + s_k \geqq \sum_{i=1}^{k} s_i.$$

But $\sum_{i=1}^{k} s_i$ is exactly the power of p that divides the left-hand side of (1). The inequality follows immediately.

THEOREM. *Let k be a positive integer. Then* $\lim_{n \to \infty} f(n)/n^k = \infty$.

Proof. Let $n \geqq (\frac{1}{2})(k+1)(k+2)^2$. Then $n/(k+1) - k/2 - 1 \geqq n/(k+2)$. Let

m be the largest integer such that $\sum_{i=0}^{k} (m+i) \leqq n$. Then

$$n < \sum_{i=0}^{k} (m+1+i) = (m+1)(k+1) + (\tfrac{1}{2})k(k+1),$$

and so

$$m > n/(k+1) - k/2 - 1 \geqq n/(k+2). \tag{2}$$

By Lemmas 1 and 2 and by inequality (2),

$$f(n) \geqq f\left(\sum_{i=0}^{k} (m+i)\right) \geqq [m, m+1, \cdots, m+k] \geqq \frac{\prod_{i=0}^{k} (m+i)}{\prod_{0 \leqq i < j \leqq k} (m+i, m+j)}$$

$$\geqq \frac{m^{k+1}}{\prod_{0 \leqq i < j \leqq k} (m+i, m+j)} > \frac{n^{k+1}}{(k+2)^{k+1} \prod_{0 \leqq i < j \leqq k} (m+i, m+j)}.$$

But $(m+i, m+j) \leqq (m+j) - (m+i) = j - i$ for $i < j$. Therefore,

$$f(n) > \frac{n^{k+1}}{(k+2)^{k+1} \prod_{0 \leqq i < j \leqq k} (j-i)}. \tag{3}$$

The theorem follows instantly from (3).

References

1. M. Hall, The Theory of Groups, Macmillan, New York, 1959.
2. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Chelsea, New York, 1953.
3. S. Shah, An inequality for the arithmetical function $g(x)$, J. Indinn Math. Soc., 3 (1939) 316–318.

NEW COMPACTIFICATIONS FROM OLD

R. E. CHANDLER, North Carolina State University

Let X, Y_1, Y_2 be topological spaces with $f_i: X \to Y_i$, $i = 1, 2$, continuous maps. The **evaluation map** $e: X \to Y_1 \times Y_2$ is defined by $e(x) = (f_1(x), f_2(x))$. A sufficient condition for e to be an embedding is that either of the f_i be an embedding [**1**, p. 78] or [**2**, p. 118]. This suggests the following construction. Let K be a compact space, $f: X \to K$ a continuous map, and cX a compactification of X. (That is, $c: X \to cX$ is an embedding and $c(X)$ is dense in cX, a compact space.) The evaluation map $e: X \to cX \times K$ as above defined, $e(x) = (c(x), f(x))$, is then an embedding of X into the compact space $cX \times K$. Let eX be the closure of $e(X)$ in $cX \times K$. Then eX is a compactification of X, generally distinct from cX. In the usual ordering of compactifications [**1**, p. 126] $eX \geqq cX$, since the restriction to eX of the projection $\pi: cX \times K \to cX$ is continuous and $\pi \circ e = c$.

Examples.

1. Let X be the countably infinite discrete space N (the positive integers), $cN = \omega N$, the one point compactification (which is homeomorphic to $\{1/n \mid n \in N\} \cup \{0\} \subseteq \mathbb{R}$, the real numbers with their usual topology), and $K = \{-1, 1\}$ with the discrete topology. Let $f: N \to K$ be defined by $f(k) = (-1)^k$. Then eN is the union of two (disjoint) convergent sequences, the two point compactification of N. We have circled the points of eN in the diagram:

$$\begin{array}{l} \odot \cdots \odot \cdot \odot \cdot \odot \quad \cdot \quad \odot \quad\quad \cdot \\ \odot \cdots \cdot \odot \cdot \odot \cdot \quad \odot \quad \cdot \quad\quad \odot \end{array}$$

$$\omega N \times K$$

Clearly, we may generalize this construction to obtain the k point compactification of N for any finite k. In a similar manner we may obtain an $\aleph_0$ point compactification by letting $K = \omega N$ and defining f to be the map

$$e(p^k) \;=\; 1/p \text{ for all primes } p,$$

and

$$e(n) \;=\; e(1) = 1 \text{ for all other integers.}$$

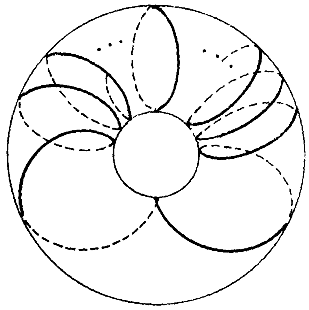
For an interesting discussion of which spaces admit finite and $\aleph_0$ point compactifications, see the two papers by Magill [3] and [4].

2. We obtain a more "exotic" compactification of N by taking for K a compact separable space with large cardinality, for example, the product of c $(= 2^{\aleph_0})$ copies of $[0, 1]$. Let D be a countable dense subset of K and define $f: N \to D \subseteq K$ so that $e^{-1}(x)$ is infinite for each $x \in D$. Again, if $cN = \omega N$ then eN turns out to be essentially $K \cup (D \times \omega N)$ with each $x \in D \subseteq K$ identified with $(x, 0) \in D \times \omega N$. (The topology, however, is such that whenever K is covered by open sets, only finitely many points of $D \times \omega N$ remain uncovered.) Thus, the cardinality of eN is 2^c. It is easily seen, though, that eN is not βN, the Stone-Cech compactification, since the map $h: N \to \{-1, 1\}$ defined by $h(x_k) = (-1)^k$, $h(n) = 1$ ($\{x_k\} = e^{-1}(x)$ for a fixed $x \in D$, n any other integer) cannot be extended to a continuous map from eN into $\{-1, 1\}$.

We can obtain βN by taking for K the product of all distinct Hausdorff compactifications of N, except ωN, $f: N \to K$ the evaluation map, and $cN = \omega N$. Then $eN = \beta N$. In fact, this is precisely the method for constructing the Stone-Čech compactification of an arbitrary completely regular space that Engelking uses [1, pp. 126–129].

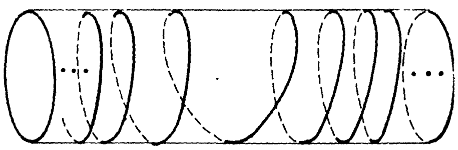
3. Let $X = \mathbb{R}$ (the real numbers with their usual topology), $c\mathbb{R} = S^1$, their one point compactification, i.e., the circle, $K = S^1$, and $f: \mathbb{R} \to S^1$ the exponential map, $f(x) = e^{2\pi i x}$; f is the mapping which wraps the line infinitely many times around the circle, once between every consecutive pair of integers. Then $e\mathbb{R}$ is the "one circle"

compactification of $\mathbb{R}$:



$c\mathbb{R} \times K$

If one uses $c\mathbb{R} = [0,1]$, the two point compactification, then $e\mathbb{R}$ is the "two circle compactification":



$c\mathbb{R} \times K$

References

1. R. Engelking, Outline of General Topology, North Holland, Amsterdam, 1968.
2. J. L. Kelley, General Topology, Van Nostrand, Princeton, N. J., 1955.
3. K. D. Magill, N-point compactifications, this MONTHLY, 72 (1965) 1075–1081.
4. ———, Countable compactifications, Canadian J. Math., 18 (1966) 616–620.

PYTHAGOREAN TRIPLES IN UNIQUE FACTORIZATION DOMAINS

K. K. KUBOTA, University of Kentucky

In two MONTHLY notes [4] and [5], Sexhauer has determined the primitive Pythagorean triples for a certain class of unique factorization domains. The aim here is to characterize Pythagorean triples in an arbitrary unique factorization domain.

Throughout this note, $D \neq (0)$ will be a unique factorization domain with field of quotients K. A Pythagorean triple in D is a triple (a,b,c) of elements of D satisfying

$$a^2 + b^2 = c^2. \tag{1}$$

It is easy to verify that if $u, v, w \in D$, then (a,b,c) and (b,a,c), where

(2) $$a = w(u^2 - v^2), \quad b = 2wuv, \quad \text{and } c = w(u^2 + v^2),$$

are Pythagorean triples in D.

Not every Pythagorean triple is of this form if D is of characteristic 2 or if 2 is neither a unit nor a prime in D. In fact, if D has characteristic 2, it is easy to see that the Pythagorean triples are those of the form $(a, b, a + b)$, where $a, b \in D$. Also, if D is a ring such that $0 \neq 2 = pq$, where $p, q \in D$ are non-units, then $(p + 2, q + 2, p + q + 2)$ is a Pythagorean triple in D. But it cannot be of the form (2) since $2 \nmid p+2$ and $2 \nmid q + 2$.

In general, if f, u, and v are arbitrary elements of D and if d is a factor of 2 relatively prime to f such that $d \mid u^2 \pm v^2$, then (a, b, c), where

(3) $$a = \frac{f(u^2 - v^2)}{d}, \quad b = \frac{2fuv}{d}, \quad \text{and } c = \frac{f(u^2 + v^2)}{d},$$

can be verified to be a Pythagorean triple. The theorem is the converse.

THEOREM. *If $D \neq (0)$ is a unique factorization domain of characteristic not 2, then every Pythagorean triple is of the form (3). If, in addition, the element 2 of D is either prime or invertible in D, then every Pythagorean triple is of the form (2).*

Proof. Let (a, b, c) be a Pythagorean triple in D. Since the case where $c - a = 0$ is trivial, we assume that $c - a \neq 0$. Then we write $c - a = gh^2$, where $g, h \in D$ and g is square-free. Define $v = h$, $u = hb/(c - a)$ and $f/d = g/2$, where $d \mid 2$ and $(f, d) = 1$. A computation using $a^2 + b^2 = c^2$ shows that these values of f, d, u, and v satisfy equation (3). It follows that $a + c = 2fu^2/d = gu^2$, so that $gu^2 \in D$. Since g is square free and $u \in K$, the field of quotients of D, it follows that $u \in D$. Also, since $(f, d) = 1$ and $a, c \in D$, equation (3) implies $d \mid u^2 \pm v^2$. Hence (a, b, c) is of the form (3).

Now suppose 2 is a unit or a prime in D. If $2 \mid g$, then $f/d = g/2 \in D$ so that (a, b, c) is of the form (2). If $2 \nmid g$, define $w = g$, $u_1 = (u + v)/2$, and $v_1 = (u - v)/2$. Then using equation (3), it is easy to see that $a = 2wu_1v_1$, $b = w(u_1^2 - v_1^2)$, and $c = w(u_1^2 + v_1^2)$. Therefore $2wu_1^2 = c + b \in D$ and $2wv_1^2 = c - b \in D$. Now $2w$ is square free since $2 \nmid w$, and $u_1, v_1 \in K$; consequently, $u_1, v_1 \in D$. Hence (b, a, c) is of the form (2) and the proof is complete.

The theorem implies that the Pythagorean triples in each of the following cases are all of the form (2):

(a) $D = Z$, the ring of ordinary integers.

(b) $D = K$, a field of characteristic not 2.

(c) $D = K[x_1, \cdots, x_n]$, where K is as in (b) or is a unique factorization domain like Z, where 2 is prime or invertible.

(d) $D = K[[x_1, \cdots, x_n]]$ (power series), where K is regular and satisfies either of the two conditions in (c).

(e) D is the ring of integers of an algebraic number field of class number 1, in which 2 is prime. For example, the cubic field of $x^3 + x + 1 = 0$.

For proofs of the facts that the rings in (c) and (d) have unique factorization, the reader is referred to Zariski and Samuel [6] and Samuel [3]. It is these two cases that motivated this work in light of Greenleaf [1], and Gross [2].

References

1. N. Greenleaf, On Fermat's equation in $C(t)$, this MONTHLY, 76 (1969) 808–809.

2. F. Gross, On the functional equation $f^n + g^n = h^n$, this MONTHLY, 73 (1966) 1093–1096.

3. P. Samuel, On unique factorization domains, Illinois J. of Math., 5 (1961) 1–17.

4. N. Sexhauer, Pythagorean triples over Gaussian domains, this MONTHLY, 73 (1966) 829–834.

5. ———, Pythagorean triples over Gaussian domains with fundamental units, this MONTHLY, 75 (1968) 278–279.

6. O. Zariski and P. Samuel, Commutative Algebra, Vol. 1, Van Nostrand, Princeton, N. J., 1958, p. 38.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

DO SELF-INTERSECTIONS CHARACTERIZE CURVES OF CONSTANT WIDTH?

B. B. PETERSON, Middlebury College

A convex curve, the boundary of a compact convex body in the Euclidean plane, has **constant width** if the distance between parallel support lines to the body is the same for all directions. On a curve of constant width w any two points at distance w lie on parallel support lines, and the chord joining them is perpendicular to the lines. Every normal to a curve of constant width is a double-normal, and this property characterizes the curves. For curves of constant width, diameters always intersect in the interior of the curve or on the curve itself. Further properties can be found in [1], [2], [4], [5], [6], [11], and [12].

For any two convex curves S_1 and S_2, we define $\alpha(S_1, S_2)$ to be the number of components of $S_1 \cap S_2$. We assume in all cases that the curves are so situated that $\alpha(S_1, S_2) > 1$, so that in particular we rule out cases where the two curves coincide or are externally tangent. In the case of two curves of constant width w, the function α can never take on odd values, although it can become infinite [10].

If it is infinite, however, the components of $S_1 \cap S_2$ can be arranged in pairs.

It follows that if a curve of constant width C intersects a congruent copy of itself C', then $\alpha(C, C')$ must be even or infinite. It has so far been impossible to find any other convex curves with this property. It is not difficult to show that convex polygons and curves with unequal chords of lateral symmetry do not have the property. Three conjectures seem worth considering:

1. If S is a convex curve and $\alpha(S, S')$ is even or infinite for every S' congruent to S, then S has constant width.

2. If S is a convex curve and $\alpha(S, C)$ is even or infinite for every circle C of diameter w, then S has constant width w.

3. If S is a convex curve and if there is a curve C of constant width w so that $\alpha(S, C')$ is even or infinite for all C' congruent to C, then S has constant width w.

Any of these statements would generalize results of Fujiwara [3], Kojima [8], Kubota [9], and Hombu [7], to the effect that if a curve can intersect itself in only two components, it is a circle.

References

1. T. Bonnesen and W. Fenchel, Theorie der konvexen Körper, Springer, Berlin, 1934; reprint Chelsea, New York, 1948.

2. H. G. Eggleston, Convexity, Cambridge, England, 1958.

3. M. Fujiwara, Ein Satz über convexe geschlossene Kurven, Sci. Rept. Tohoku Univ., 9 (1920) 289–294.

4. P. C. Hammer, Constant breadth curves in the plane, Proc. Amer. Math. Soc., 6 (1955) 333–334.

5. ———, Convex curves of constant Minkowski breadth, Proc. Symp. Pure Math. Vol. VII, Convexity, Amer. Math. Soc., (1963) 291–304.

6. ———, and A. Sobczyk, Planar line families I, Proc. Amer. Math. Soc., 4 (1953) 226–233.

7. H. Hombu, Notes on closed convex curves, Tohoku Math. J., 33 (1930) 72–77.

8. T. Kojima, On the curvature of the closed convex curve, Tohoku Math. J., 21 (1922) 15–20.

9. T. Kubota, Notes on closed convex curves, Tohoku Math. J., 21 (1922) 21–23.

10. B. Peterson, Intersection properties of curves of constant width, to appear in Illinois J. Math.

11. H. Rademacher, and O. Toeplitz, The Enjoyment of Mathematics, Princeton University Press, Princeton, N. J., 1957.

12. I. M. Yaglom and V. G. Boltyanskii, Convex Figures, Holt, Rinehart and Winston, New York, 1961.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

A TRIANGLE FOR PARTITIONS

M. O. LEVAN, Eastern Kentucky University

One of the most interesting of the number-theoretic functions is the partition function, $p(n)$, the number of ways in which the positive integer n can be expressed as a summation of positive integers. It can be explained to any bright student. The main problem beginning students run into seems to be simply computing, from the definition, enough examples to try to form a pattern.

In this note we give a "triangle" method, similar to Pascal's triangle, to reduce the amount of time involved in finding $p(n)$.

Let $p(n)$ be the number of partitions of n; $a(n)$, the number of partitions all of whose summands are odd; $b(n)$, the number all of whose summands are even; and $c(n)$, the number with at least one odd and one even summand. Clearly $p(1) = a(1)=1$; $b(1) = c(1) = 0$, and

$$p(n) = a(n) + b(n) + c(n). \tag{1}$$

Further, we have, for $n > 1$

$$b(n) = \begin{cases} p(n/2) & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd,} \end{cases} \tag{2}$$

$$c(n) = \sum_{m=1}^{n-1} a(m)b(n-m), \tag{3}$$

so that it remains to show a formula for $a(n)$.

It is well known [1] that $a(n)$ is also the number of partitions all of whose summands are distinct. We shall consider it from this viewpoint.

Let $a(n,k)$ be the number of partitions all of whose summands are distinct, and whose least summand is k. Clearly, $a(n,n) = 1$, $a(n,k) = 0$ for $n/2 \leqq k < n$ or $k > n$,

$$a(n) = \sum_{k=1}^{n} a(n,k), \tag{4}$$

and

$$a(n,k) = \sum_{j=k+1}^{n-k-1} a(n-k,j). \tag{5}$$

TABLE 1

n	$a(n)$	$a(n,1)$	$a(n,2)$	$a(n,3)$	$a(n,4)$	$a(n,5)$	$a(n,6)$	$a(n,7)$	$a(n,8)$	$a(n,9)$	$a(n,10)$	$a(n,11)$	$a(n,12)$
0	0												
1	1	1											
2	1	0	1										
3	2	1	0	1									
4	2	1	0	0	1								
5	3	1	1	0	0	1							
6	4	2	1	0	0	0	1						
7	5	2	1	1	0	0	0	1					
8	6	3	1	1	0	0	0	0	1				
9	8	3	2	1	1	0	0	0	0	1			
10	10	5	2	1	1	0	0	0	0	0	1		
11	12	5	3	1	1	1	0	0	0	0	0	1	
12													

TABLE 2

n	$a(n)$	$b(n)$	$c(n)$	$p(n)$
1	1	0	0	1
2	1	1	0	2
3	2	0	1	3
4	2	2	1	5
5	3	0	4	7
6	4	3	4	11
7	5	0	10	15
8	6	5	11	22
9	8	0	22	30
10	10	7	25	42
11	12	0	44	56
12				

We further have:

THEOREM.

$$a(n,1) = a(n-1) - a(n-1,1), \quad \textit{and} \tag{6}$$

$$a(n,k) = a(n-1,k-1) - a(n-k,k) \quad \textit{for} \quad k > 1. \tag{7}$$

Proof of the Theorem: To show (6) one may consider the partition $x_1 + \cdots + x_t = n-1$, $x_1 > \cdots > x_t$ and map it onto the partition $x_1 + \cdots + x_t + 1 = n$. This partition is counted in $a(n,1)$ unless $x_t = 1$. But the number of such partitions of $n-1$ is $a(n-1, 1)$. One may also use (4) and (5) to get

$$a(n,1) = \sum_{j=2}^{n-2} a(n-1,j) = a(n-1) - a(n-1,1).$$

Similarly, to show (7), we may consider the partition

$$x_1 + \cdots + x_t = n-1, x_1 > \cdots > x_t = k-1.$$

Then $x_1 + \cdots + (x_t + 1) = n$ is a partition of n whose least summand is k and is counted in $a(n,k)$ unless $x_{t-1} = k$. But then $x_1 + \cdots + x_{t-1} = n-k, x_1 > \cdots > x_{t-1} = k$ so that there are exactly $a(n-k,k)$ such partitions. Again, one may use (5) to get

$$\begin{aligned} a(n,k) - a(n-1,k-1) &= \sum_{j=k+1}^{n-k-1} a(n-k,j) - \sum_{j=k}^{n-k-1} a(n-1-(k-1),j) \\ &= \sum_{j=k+1}^{n-k-1} a(n-k,j) - \sum_{j=k}^{n-k-1} a(n-k,j) \\ &= -a(n-k,k). \end{aligned}$$

Using the theorem we may now construct a "triangle" for $a(n)$ and the $a(n,k)$, where it is understood any blank squares are zero. See Table 1. To compute the numbers $a(12,1)$ to $a(12,12)$ one can draw in the diagonal as shown, and from the theorem, the number in each square is the number in its upper left square minus the number on the diagonal above it. Thus

$a(12,1) = 12 - 5 = 7$; $a(12,2) = 5 - 2 = 3$; $a(12,3) = 3 - 1 = 2$;

$a(12,4) = 1 - 0 = 1$; $a(12,5) = 1 - 0 = 1$; $a(12,6) = 1 - 1 = 0$;

$a(12,7) = a(12,8) = a(12,9) = a(12,10) = a(12,11) = 0 - 0 = 0$;

$a(12,12) = 1 - 0 = 1$.

So $a(12)$ may then be computed by (4), $a(12) = 7 + 3 + 2 + 1 + 1 + 1 = 15$.

Or using (4) and (7)

$$
\begin{aligned}
a(n) &= \sum_{k=1}^{n} a(n,k) \\
&= \sum_{k=2}^{n} \{a(n-1,k-1) - a(n-k,k)\} + a(n-1) - a(n-1,1) \\
&= a(n-1) + \sum_{j=1}^{n-1} a(n-1,j) - \sum_{k=1}^{n} a(n-k,k) \\
&= 2a(n-1) - \sum_{k=1}^{n} a(n-k,k),
\end{aligned}
$$

so that $a(n)$ is twice the number above it, less the sum of the diagonal numbers. Now $a(12) = 24 - (5 + 2 + 1 + 1) = 15$.

From (3) one can then use cross multiplication of the a and b columns of Table 2 and get $c(12) = 1\cdot 0 + 1\cdot 7 + 2\cdot 0 + 2\cdot 5 + 3\cdot 0 + 4\cdot 3 + 5\cdot 0 + 6\cdot 2 + 8\cdot 0 + 10\cdot 1 + 12\cdot 0 = 51$; from the triangle, $a(12) = 15$; from (2), $b(12) = p(6) = 11$; so that from (1) $p(12) = 51 + 15 + 11 = 77$.

Reference

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford Press, London, 1960.

A COMPLETE SET WHICH IS NOT A BASIS

J. S. BYRNES, University of Massachusetts at Boston, and Naval Underwater Systems Center

Abstract: We give a straightforward example of a set which is complete, but which is not a basis, for $L^2(-\pi, \pi)$.

In elementary discussions of Fourier Series the concept of a basis is not ordinarily introduced, whereas the concept of a complete set is usually made quite clear. At subsequent levels of mathematical development the unwary might, at first, fail to recognize the vital difference between these two ideas.

In this article we give a straightforward example of a complete set which is not a basis. We note that all sequences in the paper will be defined for $-\infty < n < \infty$, and all integrals will be over the range $(-\pi, \pi)$.

We work in the space $L^2(-\pi,\pi)$ of complex-valued functions which are defined and square integrable on the interval $(-\pi,\pi)$. We say that a function $f(x) \in L^2$ is **spanned** by a sequence $\{\phi_n(x)\}$ of L^2 functions if, for any $\varepsilon > 0$, there is a finite linear combination $L_\varepsilon(x)$ of the members of the sequence (where the coefficients in this linear combination can depend upon ε), satisfying

$$\int |f(x) - L_\varepsilon(x)|^2 dx < \varepsilon.$$

If all functions in L^2 are spanned in this manner then the sequence $\{\phi_n(x)\}$ is said to be **complete** in L^2.

On the other hand, a sequence $\{\phi_n(x)\}$ is a **basis** for L^2 if, for any $f \in L^2$, there is a unique sequence $\{a_n\}$ of complex numbers satisfying:

$$\lim_{N\to\infty} \int \Big| f(x) - \sum_{n=-N}^{N} a_n\phi_n(x)\Big|^2 dx = 0.$$

We recall that the sequence $\{e^{inx}\}$ is a basis for L^2, and clearly any basis is complete.

For our example we choose the sequence defined by $\phi_n(x) = (1 + e^{ix})e^{inx}$. To show that it is complete we show that we can span each member of the sequence $\{e^{inx}\}$. Furthermore, since

$$e^{(k+1)ix} = \phi_k(x) - e^{kix} \quad \text{for } k \geqq 0 \quad \text{and} \quad e^{kix} = \phi_k(x) - e^{(k+1)ix} \quad \text{for } k < 0,$$

it is clearly sufficient to show that we can span the constant function 1. To do this we just choose a positive integer M such that $\delta = M^{-1} \leqq \varepsilon(2\pi)^{-1}$, and we take $L_\varepsilon(x) = \sum_{n=0}^{M}(-1)^n(1 - n\delta)\phi_n(x)$.

A simple calculation shows that $\int |1 - L_\varepsilon(x)|^2 dx = 2\pi M\delta^2 \leqq \varepsilon$, as required.

We now suppose that $\{a_n\}$ is a sequence of complex numbers satisfying $\lim_{M\to\infty} A_M = 0$, where

$$A_M = \frac{1}{2\pi}\int \Big| \sum_{n=-M}^{M} a_n\phi_n(x) - 1 \Big|^2 dx.$$

Setting $a_0 + a_{-1} = \alpha + \beta i$, where α and β are real, yields

$$\begin{aligned} A_M &= \frac{1}{2\pi}\int \Big| \sum_{n=-(M-1)}^{M} (a_n + a_{n-1})e^{inx} + a_{-M}e^{-iMx} + a_M e^{i(M+1)x} - 1 \Big|^2 dx \\ &= \sum_{n=-(M-1)}^{M} |a_n + a_{n-1}|^2 - (\alpha+\beta i) + |a_{-M}|^2 + |a_M|^2 - (\alpha - \beta i) + 1 \\ &= \sum_{\substack{n=-(M-1)\\ n\neq 0}}^{M} |a_n + a_{n-1}|^2 + |a_{-M}|^2 + |a_M|^2 + (\alpha-1)^2 + \beta^2 \\ &\geqq (\alpha - 1)^2 + \beta^2 \geqq 0. \end{aligned}$$

Since $A_M \to 0$ these inequalities imply that $\alpha = 1$ and $\beta = 0$, so that

$$A_M = \sum_{\substack{n=-(M-1)\\ n\neq 0}}^{M} |a_n + a_{n-1}|^2 + |a_{-M}|^2 + |a_M|^2.$$

But now the assumption that $A_M \to 0$ implies that $\lim_{M\to\pm\infty} a_M = 0$ and that $a_n + a_{n-1} = 0$ for $n \neq 0$. Thus the sequence $\{a_n\}$ must satisfy:

$$a_n + a_{n-1} = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \text{ and } \lim_{M\to\pm\infty} a_M = 0.$$

Clearly these conditions cannot be satisfied simultaneously, so that no such sequence $\{a_n\}$ exists. Thus the sequence $\{\phi_n(x)\}$ is not a basis and, as we observed previously, it is indeed complete.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

THE STIMULATION OF A MATHEMATICS STAFF—A REPORT

D. W. WESTERN, Franklin and Marshall College

Over the period of time 1968-1971 and extending into 1973, with funding provided by the National Science Foundation under two College Science Improvement Program grants, Franklin & Marshall College has had experience with different types of programs aimed at maintaining a high level of mathematical alertness on the part of the mathematics staff and at increasing their breadth of mathematical competence. This article includes a summary of that experience which, it is hoped, may be of some benefit to the mathematical community at large.

Franklin & Marshall College is an undergraduate institution with an enrollment of approximately 1900 students. The Department of Mathematics and Astronomy has a normal complement of ten of whom eight are in mathematics and one has a split load between mathematics and astronomy. The normal teaching load in mathematics is three courses per semester, a total of 12 credit hours. Student load per staff member averages about 65. The mathematics staff ranges in age from 29 to 56, seven of the eight having attained the Ph.D. with thesis topics in the fields of algebra, topology, mathematical programming, special functions, number theory, summability, and complex variables.

Participation by the Department in two separate COSIP grants has provided a continuity of program development and staff activity through three distinct stages

had by addressing the chairman of the Department of Mathematics and Astronomy, Franklin & Marshall College, Lancaster, Pennsylvania, 17604.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before August 31, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2355. *Proposed by Arthur Marshall, Madison, Wisconsin*

Given any odd integer $n > 3$, let k and j be the smallest natural numbers such that $kn + 1$ and jn are squares. Prove that n is prime if and only if both k and j exceed $n/4$.

E 2356. *Proposed by J. B. Roberts, Reed College*

If n is a natural number, define $f(n)$ to be 1 plus the sum of the prime factors of n, each prime being counted according to its multiplicity. For example, $f(12) = 8$. Prove that if n is greater than 6, then the sequence of iterates n, $f(n)$, $f(f(n)), \cdots$ contains an 8 and hence from some point on must repeat: $8, 7, 8, 7, \cdots$.

E 2357. *Proposed by M. D. Hirschhorn, Penicnik, Midlothian, Scotland*

Suppose that m and n are nonnegative integers and that $x_0, x_1, \cdots, x_m$ are distinct.

Show that

$$\sum x_0^{k_0} \cdots x_m^{k_m} = \sum_{i=0}^{m} \frac{x_i^{m+n}}{\Pi(x_i - x_j)},$$

where the sum on the left-hand side is over all $(k_0, k_1, \cdots, k_m)$ with $k_i \geqq 0$ and $k_0 + \cdots + k_m = n$, and where the product is over all $j \neq i$.

E 2358. *Proposed by W. H. Ruckle, Clemson University*

Suppose that A and B are closed convex sets and that C is bounded. Show that if $A + C = B + C$, then necessarily $A = B$.

E 2359. *Proposed by T. C. Brown, Simon Fraser University, Burnaby, Canada*

Place n distinct points on the circumference of a circle and draw all possible chords through pairs of these points. Assume no three chords are concurrent, and let a_n denote the resulting number of regions within the circle. Then the sequence $a_1, a_2, \cdots$ begins $1, 2, 4, 8, 16, 31, \cdots$. What is a_n in general?

E 2360. *Proposed by G. D. Chakerian, University of California, Davis*

A *convex body* in the plane is a convex set with non-empty interior. The *width* of a convex body is the minimum possible distance between parallel supporting lines. Show that if K is a convex body in the plane of width w and area A, then K contains a rectangle with dimensions $\sqrt{A}/4$ by $w/2$.

SOLUTIONS OF ELEMENTARY PROBLEMS

Expansion of a Symmetric Determinant

E 2297 [1971, 543]. *Proposed by Richard Stanley, Harvard University*

Let $L(n)$ be the total number of distinct monomials appearing in the expansion of the determinant of an $n \times n$ symmetric matrix $A = (a_{ij})$. For instance, $L(3) = 5$. Show that

$$\sum_{n=0}^{\infty} L(n)x^n/n! = (1-x)^{-1/2}\exp(\tfrac{1}{2}x + \tfrac{1}{4}x^2),$$

where $|x| < 1$, and where we define $L(0) = 1$.

Solution by the proposer. In J. Riordan, *An Introduction to Combinatorial Analysis*, Exercise 17, p. 44, it is stated that the coefficients of

$$A(x) = \sum_{n=0}^{\infty} a_n x^n/n! = (1-x)^{-1/2}\exp(\tfrac{1}{2}x + \tfrac{1}{4}x^2)$$

satisfy the recursion $a_{n+1} = (n+1)a_n - \binom{n}{2}a_{n-2}$. It only remains to show that $L(n+1) = (n+1)L_n - \binom{n}{2}L_{n-2}$. For a direct combinatorial proof of this, we note that $L(n)$ is equal to the number of equivalence classes in the symmetric group S_n,

where two permutations π and σ are equivalent if every cycle in the disjoint cycle decomposition of π is a cycle or the inverse of a cycle in the disjoint cycle decomposition of σ. Now for any $\pi \in S_n$, a new letter $n+1$ can be put in after any of the $1, 2, \cdots, n$ in the disjoint cycle decomposition of π, or can be left fixed. This gives $(n+1)L(n)$ new classes of permutations π. However, the two ways of inserting $n+1$ into a cycle of length 2 give the same class. There are $\binom{n}{2}$ possible cycles $(a, b) \in S_n$ of length 2, and for each one, we have counted the $L(n-2)$ classes on the set $\{1, 2, \cdots, n\} - \{a, b\}$ twice. Hence

$$L(n+1) = (n+1)L(n) - \binom{n}{2}L(n-2).$$

Also solved by Harry Lass.

Editor's Note: R. J. Dickson points out that this and similar results can be found in Pólya-Szegö, *Aufgaben und Lehrsätze II*, Berlin, 1964, pp. 310–312.

Son of E 1272

E 2298 [1971, 543, 792]. *Proposed by Anders Bager, Hjørring, Denmark*

Prove that in every triangle

$$\cos\frac{B-C}{2} + \cos\frac{C-A}{2} + \cos\frac{A-B}{2}$$

$$\leqq (\cos A + \cos B + \cos C) + \left(\sin\frac{A}{2} + \sin\frac{B}{2} + \sin\frac{C}{2}\right) \leqq 3,$$

with equality if and only if $A = B = C$.

Solution by Leon Bankoff, Los Angeles, California. By E 1272 [1957, 432; 1958, 123; 1960, 693], we have that $\sum \cos A \geqq 2 \sum \sin\frac{1}{2}B \sin\frac{1}{2}C$, with equality if and only if the triangle is equilateral. But $\sum \sin\frac{1}{2}A = \sum \cos\frac{1}{2}(B+C)$, so that

$$\sum \cos A + \sum \sin\tfrac{1}{2}A \geqq \sum \cos\tfrac{1}{2}(B+C) + 2\sum \sin\tfrac{1}{2}B \sin\tfrac{1}{2}C = \sum \cos\tfrac{1}{2}(B-C).$$

The other inequality follows immediately from (2.9) and (2.16) of O. Bottema et al., *Geometric Inequalities*, Groningen, 1969.

Also solved by V. S. Blanco, Ralph Garfield, Leonard Goldstone, M. G. Greening (Australia), Hans Kappus (Germany), Carolyn MacDonald, St. Olaf College Students, Simeon Reich (Israel), P. H. Young, and the proposer.

A Triangular Cubic

E 2299 [1971, 543]. *Proposed by Anders Bager, Hjørring, Denmark*

It is given that the roots of a certain cubic equation

$$ax^3 + bx^2 + cx + d = 0 \quad (a \neq 0)$$

are $\tan(\frac{1}{4}A)$, $\tan(\frac{1}{4}B)$, and $\tan(\frac{1}{4}C)$, where A, B, C are the angles of a triangle. Prove that $a + b = c + d$.

Solution by W. M. Sanders, Madison College, Harrisonburg, Virginia. Denote $\tan(\frac{1}{4}A)$, $\tan(\frac{1}{4}B)$, and $\tan(\frac{1}{4}C)$ by r, s, and t, respectively. Since $\frac{1}{4}A + \frac{1}{4}B + \frac{1}{4}C = \frac{1}{4}\pi$, and since $\tan(\frac{1}{4}\pi) = 1$, we obtain by elementary trigonometry

$$(*) \qquad \frac{r+s+t-rst}{1-st-rt-rs} = 1.$$

The symmetric functions of the roots of the cubic equation provide $r+s+t = -b/a$, $st+rt+rs = c/a$, and $rst = -d/a$. Substitution of these in (*) yields the desired result.

Also solved by the proposer and 77 other readers.

Editor's Comment: Arthur Boblett and C. L. Sabharwal (independently) propose the following generalization: If an n-sided planar polygon has vertex angles, $A_1, A_2, \ldots, A_n$, and if $\tan(A_1/k)$, $\tan(A_2/k), \ldots$, are the roots of the nth degree equation $a_0x^n + a_1x^{n-1} + \ldots + a_n = 0$ (where $k = 4(n-2)$), then $(a_0 + a_1) - (a_2 + a_3) + (a_4 + a_5) - + \ldots = 0$.

The Popularity of Semigroups

E 2300 [1971, 543]. *Proposed by T. C. Brown, Simon Fraser University, British Columbia*

Let S be a semigroup in which, for some fixed $k \geqq 1$, $x^{k+1} = x$ and $xy^kx = yx^ky$ for all x, y in S. Show that S is commutative.

Solution by the proposer. Since $x^{k+1} = x$, it follows that x^k is idempotent, and thus $x^2 = x^{k+2} = x(x^k)^kx = x^kx^kx^k = x^k$. Therefore $x^3 = x$ and $xy^2x = yx^2y$ for all $x, y \in S$. Hence

$$\begin{aligned} xy &= (xy)^3 = xyxyxy = xyxyx^3y = [x(yx)^2x]xy = [yxx^2yx]xy = yxyx^2y \\ &= yx(yx^2)^3y = [yx(yx^2)^2yx]xy = [yx^2(yx)^2yx^2]xy = yx^2(yx)^2yxy \\ &= yx^2(yx)^3y = yx^2yxy = yx^2yx^3y = yx[xyx^2xy] = yx[x(xy)^2x] \\ &= yx^3yxyx = (yx)^3 = yx. \end{aligned}$$

(This can be used to give a direct proof that any ring satisfying $x^3 = x$ for all x is commutative.)

Also solved by 36 other readers.

G-directed Distance Spaces

E 2301 [1971, 673]. *Proposed by David Singmaster, Bedford College, University of London, England*

Let G be a group, written additively. Define: (X, d) is a G-directed distance space if d is a function from $X \times X$ to G such that: (1) $d(x, y) = 0$ if and only if

$x = y$; (2) $d(x,y) = -d(y,x)$; (3) $d(x,z) = d(x,y) + d(y,z)$. Describe all G-directed distance spaces. (X is nonempty.)

Solution by California Polytechnic Solution Group. If $f\colon X \to G$ is one-to-one, then defining $d_f(x,y) = f(x) - f(y)$ makes (X, d_f) a G-directed distance space. Conversely, every G-directed distance space arises in this manner, for if (X,d) is such a space, choose any $x_0 \in X$ and define $f\colon X \to G$ by $f(x) = d(x, x_0)$. Then it is easy to verify that f is one-to-one and that $d = d_f$.

Also solved by twenty-four other readers and the proposer.

Editor's comment. Several solvers note that if we let $y = x$ in (3), then necessarily $d(x,x) = 0$ for all $x \in X$. Property (2) is also redundant, for we can let $z = x$ in (3) and then use the fact that $d(x,x) = 0$. A paper by the proposer on G-directed distance spaces entitled *On the concept of directed distance* has recently appeared in L'Enseignement Math. XVII, Fasc. 1 (1971), 87–91.

That's Odd, It Can Be Done

E 2302 [1971, 674]. *Proposed by Erwin Just, Bronx Community College*

Each entry a_{ij} of an nth order square matrix A is the integer $i + j \pmod{n}$. A set of n elements is selected from A so that no two elements appear in the same row, or in the same column. Prove that these n elements can be distinct if and only if n is odd.

I. *Comment by Manny Yothers, Lower Stillwater College.* The result follows immediately from E 1699 [1965, 552].

II. *Comment by Solomon Golomb, University of Southern California.* The problem can be restated as follows: Regarding the group table of the cyclic group of order n as a Latin square, when does it possess a transversal? It is well known that a group table has a transversal if and only if there is a Latin square orthogonal to it. It is also well known that the cyclic group table of order n has an orthogonal mate if and only if $n > 1$ is odd. An early reference to this is *A combinatorial problem on abelian groups*, by Marshall Hall, Jr. (Proc. AMS 3 (1952), 584–587). The definitive new reference for transversals of Latin squares, etc., is L. Mirsky, *Transversal Theory*, Academic Press, New York, 1971.

III. *Comment by C. C. Lindner, Auburn University.* It is known that a finite abelian group has a transversal if and only if it does not contain a unique element of order two. (See Marshall Hall, Jr., *op. cit.*) The matrix under consideration is the addition table for the cyclic group of order n, and the desired result is then obtained by noting that the cyclic group of order n has a unique element of order two if and only if $n > 1$ is even.

IV. *Comment by Aiden Bruen, University of Missouri, and Paul Stockmeyer, College of William and Mary.* Similar arguments will work for the case $a_{ij} = i - j$

(mod n). The latter problem is essentially solved by Martin Gardner in *Scientific American*, May, 1969, pp. 120–121.

Also solved by fifty other readers and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before August 31, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

5854. *Proposed by Stephen Gelbart, Princeton University*

Given a decreasing sequence of integers $k_1, \cdots, k_n$, a *branching* is a sequence of integers $k'_1, \cdots, k'_{n+1}$ with $k_i \geqq k'_i \geqq k_{i+1}$. Upon successively branching $n-1$ times one obtains a single integer; one calls a sequence of $n-1$ successive branchings a complete branching. Show that there are

$$\prod_{1 \leqq i < j \leqq n} [(k_i - k_j + j - i)/(j-i)]$$

distinct complete branchings of a given sequence $\{k_i\}$.

5855. *Proposed by Ioan Tomescu. Ploiesti, Rumania*

Show that any k-chromatic graph on n vertices none of which are isolated must have at least $\frac{1}{2}k(k-1) + \frac{1}{2}(n-k)$ edges.

5856. *Proposed by Jan Mycielski, University of Colorado*

For any collection X of finite subsets of a set S we denote by X^* the collection of all finite subsets T of S such that the number of subsets of T which belong to X is odd. Prove that $X^{**} = X$ and $(X\Delta Y)^* = X^*\Delta Y^*$, where $X\Delta Y = (X \cup Y) - (X \cap Y)$.

5857. *Proposed by Gérard Letac, Institut Universitaire de Technologie, Aubière, France*

$X_1, X_2, \cdots, X_t, \cdots$ being independent random variables such that $P(X_t = 0) = P(X_t = 1) = \frac{1}{2}$, define $S_t = \sum_{i=1}^{t} X_i/2^i$. Take a set H of rational numbers of the form $a/2^b$, such that H is dense in $[0,1]$. Prove or disprove that $P(\exists t > 0; S_t \in H) = 1$.

5858. *Proposed by Leonard Gallagher, University of Colorado*

Let $Q = \{r_i\}_{i=1}^{\infty}$ be any enumeration of the rationals and consider open intervals $I_i^n = N_{1/2^n}(r_i)$ about r_i. Since

$$G = \bigcap_{h=1}^{\infty} \bigcup_{i=1}^{\infty} I_i^{i+h}$$

is a G_δ set, $Q \neq G$. Demonstrate an irrational element of the G_δ set.

5859. *Proposed by L. A. Feldman, Stanislaus State College, California*

Prove that a T_0 topological space (X, T) is a metric space if and only if each $x \in X$ has a neighborhood base of open sets,

$$\{B_r(x) \mid r \in (0,1]\}$$

such that (1) if $r, s \in [0,1]$ and $B_0(x) = \{x\}$ then $B_r(x) \subset B_s(x)$; (2) if $B_r(x) \cap B_s(y) \neq \emptyset$ for $r, s \in [0,1]$, where $0 < r + s \leqq 1$, then for some t where $0 < t < r + s$, we have $x \in B_t(y)$.

SOLUTIONS OF ADVANCED PROBLEMS

The Integral of a Normalized Polynomial with Real Roots

5311 [1965, 794; 1966, 788]. *Proposed by D. Ž. Djoković, University of Waterloo, Canada*

Let $x_1 < x_2 < \cdots < x_n$ be real numbers and

$$f(x; x_1, \cdots, x_n) = (x - x_1)(x - x_2) \cdots (x - x_n),$$

$$M = \max_{x_1 < x < x_n} |f(x; x_1, \cdots, x_n)|, \quad \phi(x_1, x_2, \cdots, x_n) = \frac{1}{M} \int_{x_1}^{x_n} f(x; x_1, \cdots, x_n)\, dx.$$

Prove (or disprove) the inequality $(-1)^k(\partial\phi/\partial x_k) > 0$.

II. *Comment by Behzad Razban, Undergraduate, University of Wisconsin.* The alternate inequality, viz. $(-1)^{n+1-k}\,(\partial\phi/\partial x_k) > 0$, as proposed in the Editorial Note, is also incorrect. The basic idea is to show that $\partial\phi/dx_1 \neq 0$ when $x_1 = x_2$, and to see that this is sufficient we prove that the statement fails for $(x_1, x_2, x_3) = (c, 0, 1)$ or $(0, c, 1)$ with c small and $c < 0$ in the first case and $c > 0$ in the second. We write

$$\phi(c) = \frac{\displaystyle\int_\varepsilon^1 (x-1)x(x-c)\, dx}{M(c)},$$

where $\varepsilon = c$ in the first case and $\varepsilon = 0$ in the second and where

$$M(c) = \frac{(1-2c)(1+c)(2-c) + 2(1-c+c^2)^{3/2}}{27}.$$

Now

$$\phi'(c) = \frac{-M(c)\displaystyle\int_\varepsilon^1 (x-1)x\, dx - M'(c)\displaystyle\int_\varepsilon^1 (x-1)x(x-c)\, dx}{[M(c)]^2}.$$

$\phi'(c)$ is a continuous function of c for c small, and so $\phi'(0) = 0$ is a necessary condition for $\phi'(c)$ to have different signs for $c < 0$ and $c > 0$. But

$$\phi'(0) = \frac{-M(0)\int_0^1 (x-1)x\,dx - M'(0)\int_0^1 x^2(x-1)\,dx}{[M(0)]^2},$$

$M(0) = 4/27$, $M'(0) = -6/27$, so $\phi'(0) = 9/32 > 0$.

Note. The problem is mentioned again in Mitrinović, *Analytic Inequalities*, Springer, 1970.

Factors of $(x+2)^{2m} + x^{2m}$

5785 [1971, 305]. *Proposed by V. A. McAuley, Marshall Space Flight Center, Huntsville, Alabama*

Show that for each choice of the natural number m there are m positive numbers d_j $(j = 1, 2, \cdots, m)$ with each $d_j > 1$, such that

$$(x+2)^{2m} + x^{2m} \equiv 2\prod_{j=1}^{m}(x^2 + 2x + d_j)$$

is an identity.

Solution by R. J. Dickson, Lockheed Palo Alto Research Laboratory. A root of the polynomial on the left satisfies $|x+2| = |x|$ and hence lies on the line $\mathrm{Im}(x) = -1$. Since $x = -1$ is not a root, the roots occur in conjugate pairs with sum -2 and product exceeding unity. Since the coefficient of x^{2m} is 2, the factorization of the polynomial has the form on the right.

Also solved by the proposer and fifty-one other contributors.

Notes. A form of the problem appears as $\#$ 12 on page 221 of Durell and Robson, *Advanced Trigonometry*, London, 1936.

Many solvers applied De Moivre's theorem and obtained the values of $d_j = \csc^2[(2j-1)\pi/4m]$. David Zeitlin offers the companion identity: $(x+2)^{2m+1} + x^{2m+1} = 2(x+1)\Pi(x^2 + 2x + c_j)$.

Minimum Number of Vertices in a Four-Chromatic Graph

5786 [1971, 305]. *Proposed by Jan Mycielski, University of California, Berkeley*

Find a four-chromatic graph such that at each vertex four edges meet and each edge is contained in exactly one triangle. What is the minimum number of vertices of such a graph?

Solution by Robert Singleton, Wesleyan University. That the graph be regular of degree 4 and that each edge lie in one triangle implies the following properties:

1. There are no multiple edges or loops.
2. Each vertex lies on two of the triangles formed by the edges.

3. If triangles T_1 and T_2 have a common vertex, and so also do T_2 and T_3, then T_1 and T_3 do not have a common vertex.

Thus, the local structure of the graph is as shown by solid lines and solid circles in Figure 1. If the graph is to be minimal then it is connected.

Let G be a given graph of the type described in the problem. Let V be its set of vertices and E its set of edges. I construct an associated graph H whose sets of vertices and edges are W and F respectively. Create one vertex of H corresponding to each triangle in G. Two vertices of H are to be adjacent if and only if their corresponding triangles have a common vertex in G. Thus the edges of H correspond to the vertices of G. H, in its relation to G, may be symbolically represented by the broken lines and open circles in Figure 1. H is regular of degree 3 and, because of property (3) above, the girth of H is not less than 4.

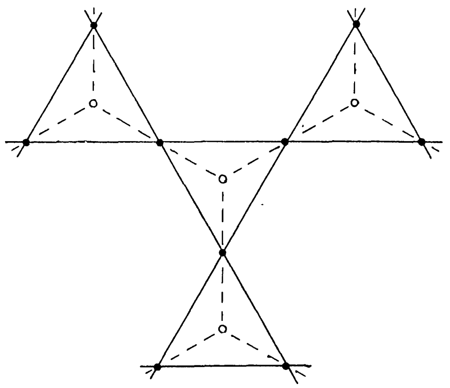


FIG. 1

Conversely, for each graph of type H one can construct a graph of type G, which is the line graph of H. Since H has no multiple edges or 3-circuits, each edge of G lies in one triangle. Since H is regular of degree 3, G is regular of degree 4.

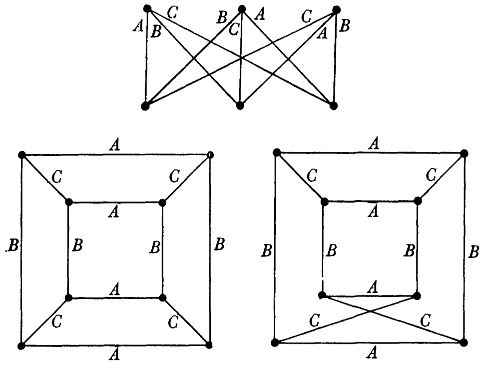


FIG. 2

Now,

$$3|W| = 2|F| = 2|V|.$$

Thus, a G of minimal order corresponds to an H of minimal order. The order of H must be even. Coloring the vertices of G corresponds to coloring the edges of H. I shall find the minimal H requiring 4 colors for edge-coloring.

The regular graphs of degree 3 and girth 4, of orders 6 and 8, are shown in Figure 2. Their edges can be colored with three colors as represented by A, B and C in the figure. The minimal regular graph of degree 3 and girth 5 is the Moore graph of order 10, shown in Figure 3. Four colors are required and the figure shows one such coloring. If one tries to use only 3 colors the outside pentagon must be colored A, B, A, B, C, starting somewhere. This determines successively the colors of the other edges and a conflict occurs.

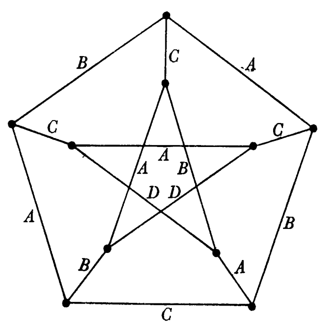


FIG. 3

Thus, a minimal graph of the type required is the line graph of the Moore graph. Its order is 15 and it has 10 triangles. It could be pictured but the picture is confused and it is better described. Take 5 triangles and stand them up around a pentagon to form a five-pointed crown. Form a similar crown with the other five triangles. Now join the points of one crown to the points of the other (that is, let their points coincide) in the following manner. Number the points of one crown 1 to 5 in sequence and those of the other 6 to 10 in sequence. Join the pairs of points: (1,6), (2,8), (3,10), (4,7), (5,9).

There are also five graphs of order 10, degree 3, girth 4, and diameter 3, but since their order is no less than that of the Moore graph I do not show them. All can be edge-colored with 3 colors.

Also solved by Frank Bernhart, Robert Connelly & Keewatin Dewdney, Don Coppersmith, Michael Doob, Jean-Paul Dufour, Vance Faber, D. P. Geller, Branko Grünbaum, Paul Himmelwright & James Williamson, A. A. Jagers (Netherlands), D. C. Kay, G. Laman (Netherlands), Sonde Nwankpa, D. P. Sumner, and J. A. Zimmer.

Non-intersecting Arcs for Nearby Points

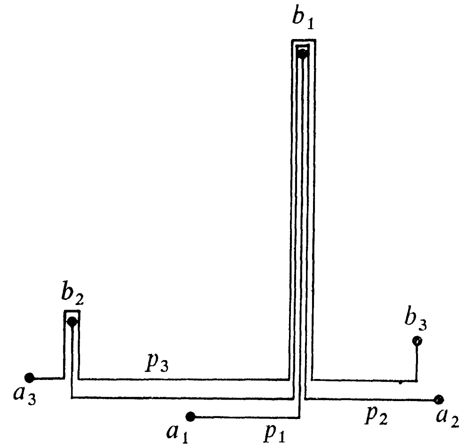
5787 [1971, 305]. *Proposed by J. L. Bryant, Florida State University*

Let $\{(a_i, b_i)\}$ be a finite collection of pairs of points in the plane each satisfying $|a_i - b_i| \leqq 1$ with all points distinct. Show that each a_i can be connected to each b_i by an arc whose diameter is no greater than $\sqrt{13}$, so that no two arcs intersect. (Diameter of an arc C means $\max(|x-y|$ for $x, y \in C)$.)

Solution by Peter Ungar, Courant Institute, New York University. We prove here the slightly stronger statement that each path can be enclosed in a square with sides < 1.

Choose Cartesian coordinate axes in such a way that neither axis is parallel to any of the $n(2n-1)$ line segments defined by the $2n$ points $a_1, \cdots, b_n$. Then the x-coordinates of all these points will be distinct and so will be their y-coordinates. Let a_i have the coordinates (a_{ix}, a_{iy}) and let b_i: (b_{ix}, b_{iy}). Let the pairs be named so that $a_{iy} < b_{iy}$ for each i. Also, let them be numbered so that $a_{1y} < a_{2y} < \cdots$.

We say a point $Q(x_0, y_0)$ is *above* an arc p if the ray $x = x_0$, $y \geqq y_0$ contains no point of p.



We construct the path p_1 from a_1 to b_1 by going horizontally from a_1 until we reach a point vertically underneath b_1 and then going straight up to b_1. Both legs of this path have length <1 and hence it is in a closed square with sides parallel to the axes and shorter than 1 and has a_1 as one of its lower corners.

We next attempt to connect a_2 to b_2 by a path p_2 in the same manner. The only obstruction to this is that the horizontal part of p_2 may have to cross the vertical part of p_1. Now the vertical part of p_1 has length <1 and it starts from a lower level than a_{2y}. Thus we can get over it by going up vertically along one side, crossing over horizontally at the top and coming down vertically to the level $y = a_{2y}$ on the other side, without getting out of a suitable square of height <1 and having a_2 as one of

its lower corners. Moreover, since the x-coordinates of all the a's and b's are different we can ensure that all the a_j and b_j with $j > 2$ will be above p_2, by just making the detour around the vertical part of p_1 sufficiently narrow.

The process can be continued without difficulty. Suppose the first $j-1$ pairs have been connected by paths consisting of horizontal and vertical segments such that a_j and b_j lie above them. To construct p_j we move from a_j horizontally towards the point (b_{jx}, a_{jy}). On the way we may have to cross peaks formed by the existing paths. These rise slightly above the height b_{iy}, where i is some index $< j$ and since $b_{iy} < a_{iy} + 1 < a_{jy} + 1$ we can pass over these peaks and still stay in a square of side 1 with a_j as one of its lower corners. Moreover, if we keep close enough to the peaks which we have to bypass then all the a_k and b_k with $k > j$ will lie above the path we are constructing, since they lie above the earlier ones.

After we reach the line $x = b_{jx}$ we complete the path p_j by going vertically up to b.

Also solved by D. J. Kleitman & Abraham Lempel.

Editorial Note. By an extensive revision of the above procedure Ungar subsequently shows that the boundary diameter $\sqrt{13}$ may not only be replaced by $\sqrt{2}$ but by $1 + \varepsilon$, ε arbitrary >0.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Modern Applied Algebra*. By Garrett Birkhoff and Thomas C. Bartee. McGraw-Hill, New York, 1970. xii + 416 pp. $11.95. (Telegraphic Review, April 1971.)

Recent years have witnessed the development of a number of interesting applications of modern algebra (see, for example, Norman Levinson's article on coding theory, this MONTHLY, March, 1970). In fact it is quite timely to have texts appearing which are devoted to the applications of modern algebra and related subjects. (The book under discussion also contains bits of graph theory, combinatorics, and

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 6

CONTENTS

(*Continued on inside cover*)

JUNE–JULY 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 11 months, Math. Notes 10 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 7 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

WHAT IS A RECIPROCITY LAW?

B. F. WYMAN, Stanford University

1. Introduction. The Law of Quadratic Reciprocity has fascinated mathematicians for over 300 years, and its generalizations and analogues occupy a central place in number theory today. Fermat's glimmerings (1640) and Gauss's proof (1796) have been distilled to an amazing abstract edifice called **class field theory**.

As a graduate student I learned the great cohomological machine and studied **Artin's Reciprocity Law**, one form of which gives an isomorphism between two cohomology groups. A little later I read Shimura's paper [**19**], called "A non-solvable reciprocity law," and couldn't understand the title at all. Where were the cohomology groups? Why was Shimura's theorem a reciprocity law?

It was an embarrassing, but healthy ignorance, because it made me go back and figure out the number theory that lay behind all those cohomology groups. Such a reassessment is especially important nowadays, because it seems more and more certain that the *next* generalization of the Law of Quadratic Reciprocity will require new techniques, and nobody is quite sure which techniques will work.

In this paper I should like to discuss reciprocity laws from a rather general but very concrete point of view. Suppose $f(X)$ is a monic irreducible polynomial with integral coefficients, and suppose p is a prime number. Reducing the coefficients of $f(X)$ modulo p gives a polynomial $f_p(X)$ with coefficients in the field $\mathbf{F}_p$ of p elements. The polynomial $f_p(X)$ may factor (even though the original $f(X)$ was irreducible). If $f_p(X)$ factors over $\mathbf{F}_p$ into a product of distinct linear factors, we say that $f(X)$ **splits completely modulo** p, and we define **Spl**(f) to be the set of all primes such that $f(X)$ splits completely modulo p.

The general **reciprocity problem** we shall be considering is: *Given $f(X)$ as above, describe the factorization of $f_p(X)$ as a function of the prime p.* Sometimes we ask for less: *give a rule to determine which primes belong to* Spl(f). This vague question is hard to make precise until it is answered. What is a "rule"? What is an acceptable method for describing the factorization of $f_p(X)$? Anyway, a satisfactory answer to this unsatisfactory question will be called a **reciprocity law**.

Quadratic polynomials are easiest to handle, and Section 2 shows how the usual Law of Quadratic Reciprocity gives a reciprocity law. (If it did not, our language would be all wrong.) Section 3 treats cyclotomic polynomials, and Sections 4 and 5 take up general results. It turns out that the reciprocity problem has been solved satisfactorily for polynomials which have an abelian Galois group, but that very little is known about polynomials whose Galois group is not abelian.

Bostwick Wyman received his Ph.D. at Berkeley in 1966 under G. Hochschild and A. Ogg. He was an Instructor at Princeton for two years and has been an Assistant Professor at Stanford since then. He spent a year's leave at the University of Oslo. His main research interest is algebraic number theory. *Editor.*

For an arbitrary polynomial $f(X)$ and a specific prime p, it only takes a finite number of steps to decide whether p is in $\mathrm{Spl}(f)$. Sections 6 and 7 give a description of an efficient algorithm for doing this calculation and report on results obtained for a family of quintic polynomials. These results probably do *not* constitute a reciprocity law, and the last section tries to answer the main question, "What is a reciprocity law?"

Prerequisites. Section 2 assumes only knowledge of the Law of Quadratic Reciprocity. The later sections assume somewhat more: acquaintance with cyclotomic polynomials, Galois groups, and the division algorithm in polynomial rings. Parts of Sections 4 and 5 assume the rudiments of algebraic number theory, but they can be skipped.

Notation. We use $\boldsymbol{Z}$, $\boldsymbol{Q}$, and $\boldsymbol{C}$ for the integers, rational numbers, and complex numbers, respectively. If q is a prime or prime power, then $\mathbf{F}_q$ is the field with q elements. If R is a ring, then $R[X]$ is the ring of polynomials with coefficients in R; mostly we deal with $\boldsymbol{Z}[X]$ and $\mathbf{F}_p[X]$.

2. Quadratic Polynomials. Suppose that $f(X)$ is an irreducible quadratic polynomial with integral coefficients. If p is a prime number, let $f_p(X)$ be the corresponding polynomial in $\mathbf{F}_p[X]$ obtained by reducing the coefficients of $f(X)$ modulo p. The reduced polynomial $f_p(X)$ can factor in one of three ways:

(0) $f_p(X) = l(X)^2$, where $l(X)$ is linear.

(1) $f_p(X) = l_1(X) \cdot l_2(X)$, where $l_1(X)$ and $l_2(X)$ are two distinct linear polynomials. In this case we say that $f(X)$ **splits** modulo p.

(2) $f_p(X)$ is irreducible in $\mathbf{F}_p[X]$.

In this paper we shall stick to polynomials of the form $X^2 - q$, where q is prime. If $f(X) = X^2 - q$, then Case (0) occurs modulo p when $p = q$, and also when $p = 2$. (The prime 2 behaves strangely for quadratic polynomials.) To distinguish Cases (1) and (2) we need to know whether q is a quadratic residue modulo p. If q is a quadratic residue, and $q \equiv a^2 \pmod{p}$, we get $X^2 - q \equiv (X + a)\,(X - a) \pmod{p}$. This puts us in Case (1) if $p \neq 2$. If q is not a quadratic residue, we are in Case (2).

Using the Legendre symbol, and ignoring the prime 2 and the exceptional Case (0) (a widespread practice!), we summarize:

(1) $X^2 - q$ splits modulo p if $(q/p) = +1$.

(2) $X^2 - q$ is irreducible modulo p if $(q/p) = -1$.

Remember that we are trying to describe the set $\mathrm{Spl}(X^2 - q)$ of primes p such that $X^2 - q$ splits modulo p, and now we know that *p is in* $\mathrm{Spl}(X^2 - q)$ *if and only if* $(q/p) = +1$.

The reader should still be skeptical, because this translation of the problem does not do much for us. The symbol (q/p) is not easy to evaluate, and besides, if we change p we have to start all over again. Since there are infinitely many primes p,

this naive approach requires an infinite amount of work to describe $\mathrm{Spl}(X^2 - q)$. Can we find a better description?

Since q is fixed and p varies, things would be better if we could use the symbol (p/q) instead of (q/p). For fixed q, the value of (p/q) depends only on the residue class of p modulo q. There are only q residue classes, and therefore only q symbols to evaluate. This suggests looking for a relationship between (p/q) and (q/p) in hopes of using (p/q) to describe $\mathrm{Spl}(X^2 - q)$. Now you can guess where we are; we have sneaked up behind the Law of Quadratic Reciprocity. Legendre's statement goes like this [**10**, p. 455 ff.]:

THEOREM 2–1 (Law of Quadratic Reciprocity): *Let p be an odd prime. Then*
1. $(1/p) = (-1)^P$, *where* $P = \frac{1}{2}(p-1)$.
2. $(2/p) = (-1)^R$, *where* $R = (p^2-1)/8$.
3. *If q is another odd prime, then* $(q/p) = (-1)^{P\cdot Q}(p/q)$, *where* $P = \frac{1}{2}(p-1)$ *and* $Q = \frac{1}{2}(q-1)$.

Gauss gave the first proof of this theorem [**6**, Article 131 ff.], and a modern proof can be found in almost any number theory text, for example, Niven and Zuckerman [**17**, p. 74].

This venerable law is really exactly what we need to compute $\mathrm{Spl}(X^2 - q)$. We start with a less fancy but quite useful form of the theorem.

THEOREM 2–2. *Let p and q be distinct odd primes.*
1. *If* $q \equiv 1 \ (mod\ 4)$, *then* $(q/p) = (p/q)$.
2. *If* $q \equiv 3 \ (mod\ 4)$, *then* $(q/p) = \begin{cases} (p/q) \text{ if } p \equiv 1 \ (mod\ 4) \\ -(p/q) \text{ if } p \equiv 3 \ (mod\ 4). \end{cases}$

The derivation of Theorem 2–2 from Theorem 2–1 is an easy exercise.

Now we are ready to give a prescription for computing (q/p) for fixed q and variable p: First, compute (b/q) for all integers b such that $1 \leqq b \leqq q-1$. Second, given p, find the b such that $1 \leqq b \leqq q-1$ and $b \equiv p \pmod q$. We have therefore $(b/q) = (p/q)$. Third, use the tables in Theorem 2 to convert knowledge of (p/q) into knowledge of (q/p).

Example 1. $q = 17$. The squares modulo 17 are 1, 2, 4, 8, 9, 13, 15, and 16, so that we have $(b/17) = +1$ for b equal one of these, and $(b/17) = -1$ for $b = 3, 5, 6, 7, 10, 11, 12$, or 14. That is (second step), $(p/17) = +1$ if and only if $p \equiv 1, 2, 4, 8, 9, 13, 15$, or 16 (mod 17). Finally, (third step), $17 \equiv 1 \pmod 4$ so that $(17/p) = (p/17)$. If we return to the language of polynomials splitting modulo a prime, we can say that

$$p \in \mathrm{Spl}(X^2 - 17) \text{ if and only if}$$

$$p \equiv 1, 2, 4, 8, 9, 13, 15, \text{ or } 16 \pmod{17}.$$

That is, the set $\mathrm{Spl}(X^2 - 17)$ can be defined by "congruence conditions modulo 17."

Example 2. $q = 11$. By finding the quadratic residues modulo 11, we conclude that $(p/11) = +1$ if and only if $p \equiv 1, 3, 4, 5$, or 9 (mod 11). In this case $11 \equiv 3$ (mod 4) so $(11/p) = \pm(p/11)$ with a sign that depends on the residue of p modulo 4. For example, $23 \equiv 1$ (mod 11), and $23 \equiv 3$ (mod 4), so that $(11/23) = -(23/11) = -(1/11) = -1$. On the other hand, $89 \equiv 1$ (mod 11) but $89 \equiv 1$ (mod 4) and $(11/89) = +(89/11) = +(1/11) = +1$. Using the Chinese Remainder Theorem, we see that the value of $(11/p)$ depends on the residue class of p modulo 44, and after some calculation we get:

$$p \in \mathrm{Spl}(X^2 - 11) \text{ if and only if}$$

$$p \equiv 1, 5, 7, 9, 19, 25, 35, 37, 39, \text{ or } 43 \pmod{44}.$$

In this case the set $\mathrm{Spl}(X^2 - 11)$ can be described by congruence conditions modulo 44.

The results of the last two examples are actually quite general.

THEOREM 2–3. *Suppose that q is an odd prime. Then the set* $\mathrm{Spl}(X^2 - q)$ *can be defined by congruence conditions modulo q if* $q \equiv 1$ (mod 4) *and modulo* $4q$ *if* $q \equiv 3$ (mod 4). *Furthermore,* $\mathrm{Spl}(X^2 - 2)$ *can be described by congruence conditions modulo* 8.

In this theorem the phrase "congruence conditions" is interpreted as in the examples. The first part follows from Theorem 2–2, and the second part from Theorem 2–1, part 2. Details are left as an exercise for the reader.

Theorem 2–3 shows that the Law of Quadratic Reciprocity gives a "reciprocity law" in the sense of Section 1. That is, it yields a nice description of sets $\mathrm{Spl}(f)$ for quadratic polynomials. In the next section we shall try to find such a reciprocity law for certain special polynomials (the *cyclotomic* ones) of higher degree.

3. Cyclotomic polynomials. Suppose ζ is a primitive nth root of unity; for instance, $\zeta = e^{2\pi i/n}$ is one choice. Then the minimal polynomial of ζ over $\boldsymbol{Q}$ is written $\Phi_n(X)$ and is called the ***n*-th cyclotomic polynomial.** One knows that $\Phi_n(X)$ has coefficients in $\boldsymbol{Z}$ and has degree $\phi(n)$, where ϕ is the Euler phi-function. It can be computed conveniently from the formula

$$X^n - 1 = \prod_{d|n} \Phi_d(X),$$

where the product runs are all divisors of n, including 1 and n itself. For example, $\Phi_1(X) = X - 1$, and if p is a prime, then $X^p - 1 = (X - 1) \cdot \Phi_p(X)$ and

$$\Phi_p(X) = X^{p-1} + X^{p-2} + \cdots + X + 1.$$

Proofs of these facts and more information about $\Phi_n(X)$ can be found in Lang [**14**, p. 206], van der Waerden [**20**, Sec. 53] and in many other algebra textbooks.

The goal of this section is a "reciprocity law" for these cyclotomic polynomials. We want a description of the set $\mathrm{Spl}(\Phi_n(X))$, and, just as in the quadratic case, the description will be given in terms of congruence conditions with respect to a modulus which depends on the polynomial. The theorem follows.

THEOREM (Cyclotomic Reciprocity Law). *The cyclotomic polynomial $\Phi_n(X)$ factors into distinct linear factors modulo p if and only if $p \equiv 1 \pmod{n}$.*

First we give a lemma about finite fields, and then use the lemma to prove the theorem. To avoid excessive notation we also use the symbol $\Phi_n(X)$ to denote the cyclotomic polynomial with coefficients reduced modulo a prime p.

LEMMA. *Suppose p is a prime number, and a is an element of $\mathbf{F}_p$ with $a^n = 1$. If $a^d \neq 1$ for all proper divisors d of n, then $X - a$ divides $\Phi_n(X)$ in $\mathbf{F}_p[X]$.*

Proof. The relation $X^n - 1 = \prod_{d|n}\Phi_d(X)$ holds in $\mathbf{F}_p$, so that $a^n - 1 = 0 = \prod_{d|n}\Phi_d(a)$. Since $\mathbf{F}_p$ is a field, it follows that $\Phi_m(a) = 0$ for some divisor m of n, and that $a^m - 1 = \prod_{d|n}\Phi_d(a) = 0$. This gives $a^m = 1$ which can only happen if $m = n$. Therefore, $\Phi_n(a) = 0$, and $X - a$ divides $\Phi_n(X)$.

Proof of theorem. Recall that the multiplicative group $\mathbf{F}_p^*$ of non-zero elements of $\mathbf{F}_p$ is cyclic of order $p - 1$. Therefore, $\mathbf{F}_p^*$ has a cyclic subgroup of order n if and only if n divides $p - 1$. Such a subgroup has $\phi(n)$ generators, so that $\mathbf{F}_p^*$ contains $\phi(n)$ distinct primitive nth roots of 1 (these generators!) if and only if it contains one, and this happens exactly when $p \equiv 1 \pmod{n}$.

Now assume $p \equiv 1 \pmod{n}$, so that $\mathbf{F}_p$ contains $\phi(n)$ distinct primitive roots of 1. These must be roots of $\Phi_n(X)$, by the lemma, so that $\Phi_n(X)$ splits into a product of distinct linear factors.

Conversely, assume that $\Phi_n(X)$ splits into linear factors modulo p. If these factors are distinct, then p cannot divide n (exercise: start from Lang [**14**, p. 206]), and it follows easily that $X^n - 1$ also has distinct roots modulo p. Let a be a root of $\Phi_n(X)$ in $\mathbf{F}_p$, so that $a^n = 1$. If d is the smallest divisor of n such that $a^d = 1$, then $\Phi_d(a) = 0$ by the lemma. If $d \neq n$, the basic relationship $X^n - 1 = \prod_{d|n}\Phi_d(X)$ shows that a is at least a double root of $X^n - 1$, a contradiction. Therefore a generates a cyclic subgroup of order n in $\mathbf{F}_p^*$, and $p \,|\, n-1$. This completes the proof of the "cyclotomic reciprocity law."

4. Abelian polynomials. In the first two sections we saw that if $f(X)$ is a quadratic or cyclotomic polynomial, then the set $\mathrm{Spl}(f)$ can be described by congruences with respect to a certain modulus. This gives a rather precise solution to the vague "reciprocity problem."

Unfortunately, such a nice description of $\mathrm{Spl}(f)$ is not always possible. We can, however, describe exactly the set of polynomials for which congruence conditions give the answer we need.

First we must recall some Galois theory. Associated to each polynomial of degree

n is the **root field** $K_f = Q(\alpha_1, \cdots, \alpha_n)$, where $\alpha_1, \cdots, \alpha_n$ are the complex roots of $f(X)$. (We avoid the more common term, "splitting field," because of possible confusion with polynomials "splitting modulo p.") The field K_f is a finite Galois extension of Q, uniquely determined by $f(x)$. The Galois group of K_f/Q is often called the **Galois group of $f(X)$**, and $f(X)$ is called an **abelian polynomial** if its Galois group is abelian.

The next theorem shows the importance this notion has for the reciprocity problem.

ABELIAN POLYNOMIAL THEOREM. *The set* $\mathrm{Spl}(f)$ *can be described by congruences with respect to a modulus depending only on* $f(X)$ *if and only if* $f(X)$ *is an abelian polynomial.*

Why should Galois groups have anything to do with polynomials splitting modulo primes? What are "congruence conditions" exactly? Enough machinery is developed in the rest of this section to establish the importance of the Galois groups, and to give a precise form of the theorem. A complete proof is far beyond the scope of this paper. In fact, the proof of the theorem involves almost all of "class field theory over the rationals." Perhaps the best avenue for an ambitious reader is to work through a basic text in algebraic number theory, and then go on to the cohomological treatment in Cassels and Fröhlich [3], or the analytic approaches of Lang [**15**], Weil [**21**], or Goldstein [7].

At this point we must escalate the prerequisites: the reader should be familiar with integral dependence, Dedekind domains, and the factorization of prime ideals in Galois extensions, or else be willing to suspend his disbelief. It is safe to skip this discussion and go on to Section 5.

Let K be an algebraic extension of Q. The elements of K whose (monic) minimal polynomial has coefficients in Z make up the ring of **algebraic integers in K**, written $\mathcal{O}_K$. The ring $\mathcal{O}_K$ is a Dedekind domain if K/Q is finite.

If p is a prime in Z, the ideal $p\mathcal{O}_K$ factors uniquely into a product of prime ideals:

$$p\mathcal{O}_K = \mathfrak{P}_1 \cdots \mathfrak{P}_r.$$

If $\mathfrak{P}$ is one of the factors of p, the residue class ring $\mathcal{O}_K/\mathfrak{P}$ is a finite field extension of Z/pZ. This **residue class field extension** is cyclic, with Galois group generated by the **Frobenius map** ϕ: $\phi(a) = a^p$ for all a in $\mathcal{O}_K/\mathfrak{P}$.

Except for a finite number of exceptions (called **ramified primes**) the $\mathfrak{P}_i$ appearing in $p\mathcal{O}_K$ are all distinct. If K/Q is Galois with group G, and p is not ramified, then for each $\mathfrak{P}_i$ there is a unique $\sigma \in G$ such that σ reduces to the Frobenius map modulo $\mathfrak{P}_i$. This automorphism is called the **Artin symbol** corresponding to $\mathfrak{P}$. We denote it by $\sigma_{\mathfrak{P}}$, so that the defining formula is

$$\sigma_{\mathfrak{P}}(a) \equiv a^p (\mathrm{mod}\, \mathfrak{P}) \text{ for all } a \in \mathcal{O}_K.$$

These Artin symbols $\sigma_{\mathfrak{P}}$ are not good enough for our purposes. We need to define an Artin symbol σ_p corresponding to a prime number p "downstairs." This is not possible in general, because different choices of the ideal $\mathfrak{P}$ may give different $\sigma_{\mathfrak{P}}$ in G. How are these various $\sigma_{\mathfrak{P}}$ related? If $\mathfrak{P}$ and $\mathfrak{Q}$ are two factors of $p\mathcal{O}_K$, then there is an automorphism τ in G such that $\tau(\mathfrak{P}) = \mathfrak{Q}$. It turns out that $\sigma_{\mathfrak{Q}} = \tau\sigma_{\mathfrak{P}}\tau^{-1}$. All the $\sigma_{\mathfrak{P}}$ corresponding to a single p are conjugate, and we call this *conjugacy class* the **Artin symbol corresponding to p**. In the good case that G is abelian, we can identify a conjugacy class with its unique member, so that the Artin symbol for p is an element σ_p in G.

EXERCISE. If you are familiar with number theory in quadratic fields, try to work out the Artin symbols for them. Start with the field $Q(\sqrt{q})$ where q is an odd prime, and identify the Galois group with $\{\pm 1\}$. Check that after this identification, the Artin symbol σ_p is exactly the Legendre symbol (q/p). (Were you wondering why σ_p is called a "symbol"?) What about more complicated quadratic fields? Finally, try to compute the Artin symbols σ_p for the cyclotomic field $Q(\zeta_m)$. (Goldstein [7, p. 96 ff.] is one of many possible references.)

From here on, K/Q is an abelian extension with group G. We denote by Q^* the multiplicative group of non-zero rational numbers, and we think of Q^* as the (multiplicative) free abelian group generated by the primes. For a fixed field K, let $\Gamma \subseteq Q^*$ be the free abelian subgroup generated by the unramified primes in K/Q. We extend the definition of the Artin symbol by setting $\sigma_{pq} = \sigma_p \cdot \sigma_q$, and $\sigma_a = \sigma_p^{-1}$ if $a = 1/p$. This procedure gives a *group homomorphism*, $\sigma: \Gamma \to G$, called the **Artin map**.

Can we find the kernel and image of this homomorphism? The image is easy to describe: *the Artin map σ is surjective.* We shall get some idea of the proof in the next section.

What about the kernel? The result here is more complicated and requires some more terminology. If a is an integer, the **ray group** Γ_a is defined as follows: a rational number $r \neq 0$ is in Γ_a if r can be written as c/d with c and d prime to a and $c \equiv d \pmod{a}$. Then *the kernel of the Artin map for K/Q contains the ray group Γ_a for some $a = p_1^{e_1} \cdots p_s^{e_s}$, where $p_1, \cdots, p_s$ are the ramified primes in K, and $e_i \geqq 1$.*

The two italicized statements above make up the **Artin Reciprocity Law**. Emil Artin conjectured it in 1923 [**1**, p. 98], and proved it in 1927 [**1**, p. 131]. (Artin worked over arbitrary number fields, not just over Q.) The theorem is central in all modern treatments of class field theory. It is proved in all the books recommended above, and in many others as well. We state it again for reference.

ARTIN RECIPROCITY LAW: *Let K/Q be a finite abelian extension with Galois group G, and let Γ be the subgroup of Q^* generated by the primes unramified in K. Then the Artin symbol gives a surjective group homomorphism σ: $\Gamma \to G$ whose kernel contains the ray group Γ_a, where a is an appropriate product of the ramified primes.*

The Artin Reciprocity Law is a precise form of half the Abelian Polynomial Theorem: if $f(X)$ is an abelian polynomial, then $\mathrm{Spl}(f)$ can be described by congruence conditions. To see why, we start with a crucial lemma.

LEMMA. *Suppose $f(X)$ is an abelian polynomial with root field K, Galois group G, and Artin map $\sigma:\Gamma\to G$. Then except perhaps for a finite number of exceptional primes, $f(X)$ splits modulo p if and only if σ_p is trivial.*

Proof. We can only give an outline here. If p is unramified and $p\mathcal{O}_K=\mathfrak{P}_1\cdots\mathfrak{P}_s$, then the Chinese Remainder Theorem gives

$$\mathcal{O}_K/p\mathcal{O}_K\cong\oplus_{i=1}^s\mathcal{O}_K/\mathfrak{P}_i.$$

On the other hand, except for a finite number of p,

$$\mathcal{O}_K/p\mathcal{O}_K\cong\mathbf{F}_p[X]/(f_p(X)),$$

where $f_p(X)$ is the reduction of $f(X)$ modulo p. (This is a hard exercise; the exceptions all divide the discriminant of $f(X)$.) Therefore, except for finitely many p,

$$\mathbf{F}_p[X]/(f_p(X))\cong\oplus_{i=1}^s\mathcal{O}_K/\mathfrak{P}_i.$$

When is the Artin symbol σ_p trivial? By definition σ_p induces the Frobenius map, $x\to x^p$, on each direct summand. The Frobenius map is trivial on $\mathcal{O}_K/\mathfrak{P}$ only if $\mathcal{O}_K/\mathfrak{P}\cong\mathbf{F}_p$, so that $\mathbf{F}_p[X]/(f_p(X))\cong\mathbf{F}_p^n$ when σ_p is trivial, and this is only possible when $f_p(X)$ factors into linear factors. All the steps are reversible, so the converse holds too.

This lemma, combined with the Artin Reciprocity Law, guarantees that the set $\mathrm{Spl}(f)$ contains all primes p such that $p\equiv 1\pmod a$, with at most finitely many exceptions. (Check this!)

We need to change $\mathrm{Spl}(f)$ slightly at this point. Add to $\mathrm{Spl}(f)$ any primes $p\equiv 1$ (mod a) not already there, and throw away any divisors of a. Call the resulting set S; this is the set we can describe by explicit congruence conditions.

Let $\boldsymbol{Q}^*(a)$ be the multiplicative subgroup generated by all primes p which do not divide a. (A fraction b/c in lowest terms is in $\boldsymbol{Q}^*(a)$ if both b and c are prime to a.) Let S' be the subgroup of $\boldsymbol{Q}^*$ generated by S. The set S has been chosen so that $\Gamma_a\subseteq S'\subseteq\boldsymbol{Q}^*$, and the importance of these inclusions comes out in the next lemma.

LEMMA. *$\boldsymbol{Q}^*(a)/\Gamma_a\cong(\boldsymbol{Z}/a\boldsymbol{Z})^*$, where $(\boldsymbol{Z}/a\boldsymbol{Z})^*$ is the group of invertible elements in $\boldsymbol{Z}/a\boldsymbol{Z}$.*

Proof. Define $\theta:\boldsymbol{Q}^*(a)/\Gamma_a\to(\boldsymbol{Z}/a\boldsymbol{Z})^*$ by $\theta(b/c)=bc^{-1}$ (mod a). Check as an exercise that θ is a surjective homomorphism with kernel exactly Γ_a.

This lemma supplies us with congruence conditions. Starting with $\mathrm{Spl}(f)$, pass to S, and consider the set $\theta(S')$ of residue classes modulo a. A given prime p will lie in S if and only if its residue class modulo a lies in $\theta(S')$. Since S and $\mathrm{Spl}(f)$ differ in at most a finite number of primes, we shall be content with this result.

Next we attack the other half of the Abelian Polynomial Theorem: If $\mathrm{Spl}(f)$ can be defined by congruences, then $f(X)$ must be an abelian polynomial. We shall need a hard theorem which says (roughly) that the root field $\boldsymbol{K}_f$ of $f(X)$ is uniquely determined by the set $\mathrm{Spl}(f)$. We introduce some notation: If S and T are two sets of primes, then $S \subseteq^* T$ means that except for at most a finite number of exceptions every member of S is a member of T. The precise statement is then:

INCLUSION THEOREM. *Suppose $f(X)$ and $g(X)$ are polynomials with root fields $\boldsymbol{K}_f$ and $\boldsymbol{K}_g$, respectively. Then $\boldsymbol{K}_f \subseteq \boldsymbol{K}_g$ if and only if* $\mathrm{Spl}(g) \subseteq^* \mathrm{Spl}(f)$.

Note the reversal! The similarity to Galois theory can be made very precise for abelian polynomials and is an important part of class field theory. The theorem itself holds for arbitrary $f(X)$ and $g(X)$.

It is not hard to prove that $\boldsymbol{K}_f \subseteq \boldsymbol{K}_g$ implies $\mathrm{Spl}(g) \subseteq^* \mathrm{Spl}(f)$. The converse requires analytic techniques, and is a corollary of the Tchebotarev Density Theorem discussed in the next section. See Cassels and Fröhlich [3, Exercise 6.1, p. 362] or Goldstein [7, Theorem 9–1–13, p. 164] for a proof.

Assume now that $\mathrm{Spl}(f)$ can be defined by congruences modulo an integer a. Actually we assume more: namely, that $\mathrm{Spl}(f)$ contains the ray group Γ_a. (Exercise: What's the difference between these assumptions?) According to Section 3, Γ_a is $\mathrm{Spl}(\Phi_a(X))$, and the root field of $\Phi_a(X)$ is the cyclotomic field $\boldsymbol{Q}(\zeta_a)$, which is abelian over $\boldsymbol{Q}$. Since $\Gamma_a \subseteq \mathrm{Spl}(f)$, the Inclusion Theorem gives $\boldsymbol{K}_f \subseteq \boldsymbol{Q}(\zeta_a)$, so that $\boldsymbol{K}_f$ must also be abelian over $\boldsymbol{Q}$.

One corollary of this discussion deserves special mention.

KRONECKER'S THEOREM. *Every abelian extension of $\boldsymbol{Q}$ is contained in a cyclotomic extension.*

Proof. Exercise: Combine the Artin Reciprocity Law with the argument above. (There is an elementary proof in Gaal [5, p. 242].)

5. General polynomials. The Tchebotarev Density Theorem. If $f(X)$ is an irreducible polynomial in $\boldsymbol{Z}[X]$ which is not abelian, then very little can be said about the set $\mathrm{Spl}(f)$. The best general result is a statement about the relative "size" of $\mathrm{Spl}(f)$. First we describe a numerical measure of sets of primes called the **density**.

Let Π be the set of all prime numbers, and let $T \subseteq \Pi$ be any subset. For any real $x \geqq 1$, let

$$\delta(T,x) = \frac{\mathrm{card}\{p \in T \mid p < x\}}{\mathrm{card}\{p \in \Pi \mid p < x\}}.$$

DEFINITION. *If T is a set of primes such that* $\lim_{x\to\infty} \delta(x,T) = \delta(T)$, *then T has density* $\delta(T)$.

Note that the limit may not exist. In that case we say, naturally enough, that T "does not have a density." If T does have a density, then $0 \leqq \delta(T) \leqq 1$. Since Π is infinite, any finite set of primes has density 0, and it is easy to see that if S and T differ by a finite set of primes, then $\delta(S) = \delta(T)$. Clearly $\delta(\Pi) = 1$.

One can prove that a set of primes is infinite by showing that it has a non-zero density. The first theorem of this type was proved by Lejeune Dirichlet in 1837.

DIRICHLET'S THEOREM. *Suppose m is a positive integer and a is an integer relatively prime to m. Then the set of all primes congruent to a modulo m has a density equal to $1/\phi(m)$.*

In particular, the set of all primes congruent to a modulo m is infinite. Although this much can be proved directly for some a and m (see Hardy and Wright [**8**, p. 13]), no general proof avoids analysis and the notion of density. Proofs of the theorem can be found all over; one is in Davenport [**4**, pp. 1 and 28].

The density result we need for the reciprocity problem is the *Tchebotarev Density Theorem.* We give a weakened version first.

WEAK TCHEBOTAREV THEOREM. *Let $f(X)$ be an irreducible polynomial in $\mathbf{Z}[X]$ with root field $\mathbf{K}_f$, and suppose that $[\mathbf{K}_f : \mathbf{Q}] = n$. Then* Spl$(f)$ *has a density equal to $1/n$.*

This theorem implies part of Dirichlet's Theorem. Take $f(X)$ to be the cyclotomic polynomial $\Phi_m(X)$ so that $[\mathbf{K}_f : \mathbf{Q}] = \phi(m)$, so that the theorem gives $\delta[\mathrm{Spl}(\Phi_m)] = 1/\phi(m)$. By Section 3, a prime p is in $\mathrm{Spl}(\Phi_m)$ if and only if $p \equiv 1 \pmod{m}$ and putting all this together gives Dirichlet's result for $a = 1$. The rest of Dirichlet's Theorem follows from the full Tchebotarev Theorem discussed below.

The interested reader should go back to Section 2 and examine quadratic polynomials from the point of view of density results. The following main result can be derived from either of the two theorems above: *Suppose a is not a perfect square. Then the set of primes p such that $(a/p) = +1$ has density $\frac{1}{2}$.* (What about those p with $(a/p) = -1$? What about primes dividing a?)

To explain the strong form of Tchebotarev's theorem, we need to use Artin symbols again. To read the rest of this section you need either the last part of Section 4 or faith. It is safe to skip to Section 6.

Let $f(X)$ in $\mathbf{Z}[X]$ have root field $\mathbf{K}_f$ and Galois group G. The group G is not necessarily abelian, and the Artin symbol corresponding to p is a conjugacy class C_p of elements of G. (There are a finite number of ramified p for which C_p cannot be defined. We ignore these.)

Tchebotarev proved his theorem in 1925 and his methods inspired Artin's proof of the Reciprocity Law.

TCHEBOTAREV DENSITY THEOREM. *Let $f(X) \in Z[X]$ be irreducible with Galois group G, and let C be a fixed conjugacy class of elements of G. Let S be the set of*

primes p whose Artin symbol C_p equals C. Then S has a density, and

$$\delta(S) = \frac{\operatorname{card}(C)}{\operatorname{card}(G)}.$$

In particular, if $C = \{1\}$, then $S = \operatorname{Spl}(f)$ (by a lemma in Section 4) and $\delta(S) = 1/\operatorname{card}(G)$. We recover the weak theorem. If the group G is abelian, then each conjugacy class has one member and the corresponding sets of primes each have density $1/\operatorname{card}(G)$. This shows immediately that *the Artin map is surjective.* (Why?) Also explicit calculation of Artin symbols in cyclotomic fields gives a proof of Dirichlet's Theorem from Tchebotarev's Theorem.

6. An algorithm for the reciprocity problem. What have we learned so far about the reciprocity problem? Not much, in general, but we can claim to understand abelian polynomials completely. This knowledge at least gives a starting place for the study of polynomials with *solvable* Galois group. We do not discuss this here, but see Hasse [**9**, pp. 64–69] and Cassels and Fröhlich [**3**, Ex. 2.15, p. 354]. For polynomials with non-solvable groups, the only progress is the tantalizing example of Shimura mentioned in the introduction.

No satisfactory description of general sets $\operatorname{Spl}(f)$ has been given up to now, but for fixed $f(X)$ and a particular prime p, we can at least ask whether p lies in $\operatorname{Spl}(f)$. This involves factoring $f(X)$ modulo p, which is a finite process. The point of this section is to do the factoring *efficiently*. The method we use is essentially due to Berlekamp [**12**, Chapter 6]. Our formulation, designed to give only that information relevant to the reciprocity problem, is slightly different from Berlekamp's.

The prerequisites for the discussion are the Chinese Remainder Theorem for polynomial rings, and some knowledge of finite fields. (The material needed is covered in Berlekamp [**2**] and Lang [**14**], especially pages 63 and 182.)

Suppose given a polynomial $f(X)$ in $\mathbf{Z}[X]$ of degree n, with no repeated factors, and let $f_p(X)$ be its reduction modulo p. Assume $f_p(X) = g_1(X) \cdots g_r(X)$ where $g_i(X)$ is irreducible of degree d_i. Our problem is to compute $d_1, \cdots, d_r$; we know $d_1 + \cdots + d_r = n$. For example, $p \in \operatorname{Spl}(f)$ if $r = n$ and each $d_i = 1$.

First we compute the discriminant $D(f)$ by the classical formula (e.g., Lang [**14**, p. 139]). If p divides $D(f)$, then $f_p(X)$ has a repeated factor. We declare such p "bad" and do not consider them further. If p does not divide $D(f)$, then the $g_i(X)$ are distinct irreducible polynomials and are therefore relatively prime. The Chinese Remainder Theorem gives:

$$(*) \qquad \mathbf{F}_p[X]/(f_p(X)) \cong \oplus_{i=1}^r \mathbf{F}_p[X]/(g_i(X)).$$

We write $\mathbf{A} = \mathbf{F}_p[X]/(f_p(X))$ and $k_i = \mathbf{F}_p[X]/(g_i(X))$. Since $g_i(X)$ is irreducible of degree d, then $k_i = \mathbf{F}_q$, the unique finite field with $q = p^{d_i}$ elements. Since $[k_i : \mathbf{F}_p] = d_i$, we can recover all we need by computing the dimensions of the summands on the righthand side.

Here we have a case in which two isomorphic structures cannot be identified: the ring Λ is given very concretely as an n-dimensional $\mathbf{F}_p$ space, with basis $1, x^2, \cdots, x^{n-1}$, where x is the residue class of X modulo $f(X)$. Addition is vector space addition, and multiplication is carried out modulo $f(X)$. Our problem is to extract the direct sum decomposition, or at least compute the d_i, from *this* description of Λ.

As preparation, consider a finite extension $\boldsymbol{k}$ of $\mathbf{F}_p$, with $[\boldsymbol{k}:\mathbf{F}_p] = d$, say. The mapping $\phi(z) = z^p: \boldsymbol{k} \to \boldsymbol{k}$ is a field isomorphism called the Frobenius map, and $\phi(z) = z$ if and only if $z \in \mathbf{F}_p$. Moreover, $\phi^i(z) = z$ for $1 \leqq i \leqq d$ if and only if $z \in \mathbf{F}_q \subseteq \boldsymbol{k}$, where $q = p^i$. Thus, d can be computed as the smallest integer such that $\phi^d =$ identity on $\boldsymbol{k}$.

The Frobenius map $z \to z^p$ on Λ, which we also denote by ϕ, is a ring isomorphism useful in studying the structure of Λ. For example, if $\Lambda \cong \mathbf{F}_p \oplus \cdots \oplus \mathbf{F}_p$ (n summands), then $\phi =$ identity. More generally, the smallest d such that $\phi^d =$ identity on Λ (the order of ϕ) equals the least common multiple of the d_i. Since x generates Λ as a ring, the order is the smallest d such that $\phi^d(x) = x$, so it is easy to compute. We shall see in the next section that the order can give a lot of information in special cases. In general, however, we need a refinement.

Suppose γ denotes the isomorphism in the Chinese Remainder Theorem:

$$\gamma: \Lambda \cong \boldsymbol{k}_1 \oplus \cdots \oplus \boldsymbol{k}_r.$$

Then it is easy to see that

$$\gamma(\ker(\phi - I)) = \mathbf{F}_p \oplus \cdots \oplus \mathbf{F}_p, \ r \text{ summands},$$

where $I: \Lambda \to \Lambda$ is the identity map, and $\ker(\phi - I)$ is the kernel of the linear transformation $(\phi - I): \Lambda \to \Lambda$.

Similarly,

$$\gamma(\ker(\phi^2 - I)) = \boldsymbol{l}_1 \oplus \cdots \oplus \boldsymbol{l}_r,$$

where $\boldsymbol{l}_i = \mathbf{F}_{p^2}$ if $\mathbf{F}_{p^2} \subseteq \boldsymbol{k}_i$, and $\boldsymbol{l}_i = \mathbf{F}_p = \boldsymbol{k}_i$, otherwise.

Therefore, $\ker(\phi^2 - I)$ has $\mathbf{F}_p$-dimension equal to $2r -$(the number of summands with $d_i = 1$).

DEFINITION. *For each integer i, let* $\nu_i = \text{nullity } (\phi^i - I) = \dim(\ker(\phi^i - I))$, *where "dim" denotes vector space dimension over the prime field* $\mathbf{F}_p$.

For each integer j, let $\mu_j =$ *the number of factors in the decomposition* (*) *which have dimension exactly equal to j.*

In this notation $\nu_1 = r$, the total number of factors, and $\nu_2 = 2r - \mu_1$. The reader should verify that

$$\begin{aligned} \nu_3 &= \mu_1 + 2\mu_2 + 3(r - \mu_1 - \mu_2) \\ &= 3r - 2\mu_1 - \mu_2. \end{aligned}$$

Generally, it is not hard to see that

$$(\#) \qquad \nu_k = kr - (k-1)\mu_1 - (k-2)\mu_2 - \cdots - \mu_{k-1}.$$

This relationship is very important. Knowing the μ_i is the same as knowing $d_1, d_2, \cdots, d_r$, so they give the factorization of $f_p(X)$. On the other hand, we shall see below that the ν_i are relatively easy to compute. The reader should use equation (#) to verify the following inversion formula:

$$(\#\#) \qquad \mu_k = 2\nu_k - \nu_{k-1} - \nu_{k+1}.$$

We summarize these facts in the theorem.

THEOREM. *Suppose, given* $\Lambda = \mathbf{F}_p[X]/(f_p(X)) = \boldsymbol{k}_1 \oplus \cdots \oplus \boldsymbol{k}_r$, *and let* $d_i = [\boldsymbol{k}_i : \mathbf{F}_p]$. *Let* ϕ *be the Frobenius automorphism of* Λ, *and let* $\nu_i =$ *nullity* $(\phi^i - I)$. *Then* $r = \nu_1$, *and there are exactly* $\mu_j = 2\nu_j - \nu_{j-1} - \nu_{j+1}$ *summands with* $d_i = j$, $j = 1, \cdots, d$. *Here* d *is the smallest integer such that* $\phi^d = I$.

This theorem forms the basis of an efficient algorithm. First compute the matrix $[\phi]$ with respect to the basis $\{1, x, \cdots, x^{n-1}\}$ of Λ. (Berlekamp calls this the **Q-matrix.**) Then compute successively $\nu_i =$ nullity $([\phi]^i - I)$. Finally, compute the μ_i from the theorem. If $\mu_1 = n$, then p belongs to $\mathrm{Spl}(f)$, and in more complicated situations the μ_i give information about the Artin symbol belonging to p.

Of course, we must examine this proposed algorithm. How hard is it? How long does it take? Can it produce significant results and lead to a better theoretical understanding of the problem?

First of all I have to admit that it is completely unreasonable to do the algorithm by hand. I worked on $f(X) = X^5 - X - 1$ with $p = 11$ for an hour and could not make it come out. It is much easier to factor by trial and error when p is small, but large primes are impossible.

Fortunately it is not too difficult to write a FORTRAN program which will do calculations in the ring Λ. Since Λ is an n-dimensional vector space over $\mathbf{F}_p$ with a nice basis $\{1, x, x^2, \cdots, x^{n-1}\}$, its elements can be represented as a $1 \times n$ FORTRAN array. The program written for the next section uses FORTRAN's integer arithmetic and works modulo a variable prime p.

The algorithm is very efficient in that *the number of operations required to factor* $f(X)$ *modulo* p *is proportional to* $\log p$. In fact, the only part of the algorithm that depends essentially on p is computing x^p in the ring Λ. Abstractly speaking, how many steps does it take to compute x^p? Certainly less than $2 \cdot \log_2 p$, since x^p can be computed by successively squaring together some multiplications by x. (Are you skeptical? If $p = 23 = 10111$ (binary), the steps are x, x^2, x^4, x^5, x^{10}, x^{11}, x^{22}, x^{23}, which requires $7 < 2 \cdot \log_2(23)$ steps.) The fascinating subject of number theory algorithms and the time needed to do them is discussed in Lehmer [**16**]. Knuth [**13**, p. 388 ff.] goes into more detail and discusses algorithms very similar to this one.

7. Numerical results. With the help of R. W. Latzer I have written a FORTRAN program to carry out the algorithm for the polynomials $X^5 - X - a$, where a is an integer. This is the "Bring-Jerrard Quintic" which has the non-solvable Galois group $\mathfrak{S}(5)$ for general a, and in particular for $a = 1$, and $a = 2$. The program factored $X^5 - X - 1$ for all p up to 23,099 in about two minutes, at which time the program overflowed the FORTRAN integer capacity. (I have learned that Professor J. D. Brillhart, using other methods, has factored many members of a more general family of quintics up to $p = 1000$.)

If $f(X)$ is any irreducible quintic polynomial, then $f_p(X)$ can factor in one of eight ways:

Type 0:	$p \mid D(f)$	
Type 1:	Five linear factors	1/120
Type 23:	(Quadratic) (Quadratic) (Linear)	15/120
Type 24:	(Quadratic) (Three Linear)	10/120
Type 3:	(Cubic) (Linear) (Linear)	20/120
Type 4:	(Quartic) (Linear)	30/120
Type 5:	(Quintic)	24/120
Type 6:	(Quadratic) (Cubic)	20/120

The factors are irreducible and distinct, when displayed. The type is the order d of the Frobenius map when p does not divide $D(f)$ except that Type 23 means (order $= 2$, nullity $v_1 = 3$) and Type 24 means (order $= 2$, nullity $v_1 = 4$). Thus, no nullities have to be computed, except when the order $= 2$. The fractions give the density of primes of each type, according to the Tchebotarev Density Theorem.

Finally, we give some examples of actual numerical results.

1. $f(X) = X^5 - X - 1$.

(a) $D(f) = 19 \cdot 151$, so 19 and 151 are bad.

(b) The primes of Type 1 (those in $\mathrm{Spl}(f)$) which are less than 23099 are 1973, 3769, 5101, 7727, 8161, 9631, 11093, 14629, 16903, 17737, 17921, 18097, 19477, 20759, 21727, and 22717. There are 16 primes in this list, giving a ratio of 16/2350 $\approx .0068$, as compared with a density of $1/120 \approx .00833$.

(c) The primes less than 500 are classified as follows:

Type 0. 19, 151.
Type 1. None.
Type 23. 67, 71, 239, 251, 313, 421, 433, and 491.
Type 24. 163, 193, 227, 307, 467, 487, and 499.
Type 3. 17, 41, 43, 47, 53, 107, 113, 179, 181, 191, 229, 281, 293, 311, 317, 347, 349, 373, 409, 457, and 463.

Type 4. 23, 29, 31, 61, 97, 101, 127, 131, 157, 173, 223, 241, 263, 269, 331, 359, 389, 439, 443, and 479.

Type 5. 3, 5, 11, 13, 79, 89, 109, 137, 139, 211, 257, 337, 379, 397, 431, 449, and 461.

Type 6. 2, 7, 37, 59, 73, 83, 103, 149, 167, 197, 199, 233, 271, 277, 283, 353, 367, 383, 401, and 419.

2. $f(X) = X^5 - X - 2$.

(a) $D(f) = 2^4 \cdot 3109$, so 2 and 3109 are bad.

(b) The primes of type 1 less than 23099 are 229, 271, 1637, 2647, 2857, 3673, 6323, 7103, 8123, 8999, 11161, 12197, 14341, 14503, 14929, 17183, 18679, 19457 and 20563. There are 19 primes in this list, giving $19/2350 \approx .00809$.

3. It is also possible to fix p and let the coefficient a in $X^5 - X - a$ vary modulo p. So far I have done this for all p up to 239. For example, if $p = 31$, we get:

Type 0. $a \equiv 11, 20 \pmod{31}$.

Type 1. None.

Type 23. $a \equiv 2, 3, 28, 29 \pmod{31}$.

Type 24. $a \equiv 0, 15, 16 \pmod{31}$.

Type 3. $a \equiv 7, 24 \pmod{31}$.

Type 4. $a \equiv 1, 5, 8, 9, 14, 17, 22, 23, 26, 30 \pmod{31}$.

Type 5. $a \equiv 6, 10, 12, 13, 18, 19, 21, 25 \pmod{31}$.

Type 6. $a \equiv 4, 27 \pmod{31}$.

8. What is a reciprocity law? A general reciprocity law should provide a description of the set $\mathrm{Spl}(f)$ associated with a polynomial $f(X)$. The algorithm discussed in this paper is such a description, but few number theorists would consider it a reciprocity law. More is wanted, but the exact requirements are still vague and undefined.

A good general reciprocity law should specialize to the Artin Reciprocity Law in the case of abelian polynomials. A very good reciprocity law should include a one-to-one correspondence between certain sets of prime numbers and field extensions, giving more substance to the Inclusion Theorem in Section 5. Such a correspondence should generalize the known abelian theorems of class field theory. Y. Ihara [**11**] is beginning to make some progress toward this goal in the function field case.

Even if a good correspondence cannot be set up, any reciprocity law must be set in a general framework, and should unify various kinds of number theoretic phenomena. The examples in Shimura [**19**] are related to the theory of elliptic curves, but they are very special, and it is not clear how to use them as a foundation for a general reciprocity law. (The specialist should look at Ihara's discussion of this question.)

I would like to mention briefly another direction of research which may lead to reciprocity laws. The Artin Reciprocity Law can be interpreted as a theorem about certain classes of analytic functions: see Artin's original paper [**1**, p. 97] or the section "Abelian L-functions are Hecke L-functions" in Goldstein [**7**, p. 182]. There seem to be important non-abelian analogues to this viewpoint which involve group

representations and automorphic forms, and the interested reader should look at the introduction to Jacquet-Langlands' book [12] or Shalika's paper [18].

Finally, I have to confess that I still do not know what a reciprocity law is, or what one should be. The reciprocity problem, like so many other number theory problems, can be stated in a fairly simple and concrete way. However, the simply stated problems are often the hardest, and a complete solution seems to be far out of reach. In fact, we probably will not know what we are looking for until we have found it.

This research was supported in part by a grant from the National Science Foundation under grant GP 29696.

References

1. E. Artin, Collected Papers, Addison-Wesley, Reading, Mass., 1965.
2. E. R. Berlekamp, Algebraic Coding Theory, McGraw-Hill, New York, 1968.
3. J. W. S. Cassels and A. Fröhlich, Algebraic Number Theory, Thompson, Washington, 1967.
4. H. Davenport, Multiplicative Number Theory, Markham, Chicago, 1971.
5. L. Gaal, Classical Galois Theory with Examples, Markham, Chicago, 1971.
6. C. F. Gauss, Disquisitiones Arithmeticae, transl. A. A. Clarke, Yale, New Haven, 1966.
7. L. Goldstein, Analytic Number Theory, Prentice-Hall, Englewood Cliffs, N. J., 1971.
8. G. Hardy and E. Wright, An Introduction to the Theory of Numbers, 4th ed. Oxford, 1960.
9. H. Hasse, Bericht über neuere Untersuchungen und Problemen aus der Theorie der algebraischen Zahlkörper, Teil I, Ia, II, 2nd edit., Physica-Verlag, Würzburg, 1965.
10. A. M. Legendre, Recherches d'Analyse Indéterminée, Hist. Acad., Paris, 1785.
11. Y. Ihara, Non-abelian class fields over function fields in special cases, to appear in Proc. of the Intern. Congress of Math., Nice, 1970.
12. H. Jacquet and R. P. Langlands, Automorphic Forms on Gl(2), Springer-Verlag Lecture Notes in Mathematics, No. 114, Berlin, 1970.
13. D. Knuth, The Art of Computer Programming, Volume 2: Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1969.
14. S. Lang, Algebra, Addison-Wesley, Reading, Mass., 1965.
15. ———, Algebraic Number Theory, Addison-Wesley, Reading, Mass., 1971.
16. D. H. Lehmer, Computer Technology Applied to the Theory of Numbers, Studies in Number Theory, MAA, Prentice-Hall, Englewood Cliffs, N.J., 1969.
17. I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, 2nd ed., Wiley, New York, 1966.
18. J. Shalika, Some Conjectures in Class Field Theory, in AMS, Proc. of Symposia in Pure Math., Volume XX: Stony Brook Number Theory Institute, Providence, 1971.
19. G. Shimura, A non-solvable reciprocity law, J. Reine Angew. Math., 221 (1966) 209–220.
20. B. van der Waerden, Modern Algebra, Vol. I, Revised English Edition, Ungar, New York, 1953.
21. A. Weil, Basic Number Theory, Springer, New York, 1967.

A MAP OF SOURCES, SINKS AND SADDLES

D. M. JORDAN AND H. L. PORTEOUS, University of Hull

The system of linear differential equations

$$\dot{x}_1 = ax_1 + bx_2, \qquad \dot{x}_2 = cx_1 + dx_2$$

yields a plane flow which has been traditionally classified into one of a number of types including source, sink, saddle, spiral, centre, and node. We define a geometrical equivalence relation which gives this classification, the equivalence classes being called directed-orbit-types. Also we give a map of this classification using a topology which makes the set of these flows homeomorphic to $\mathbf{R}^4$.

1. Classifications. The system

$$\dot{x}_1 = ax_1 + bx_2, \qquad \dot{x}_2 = cx_1 + dx_2,$$

where $a, b, c, d \in \mathbf{R}$, gives rise to the plane flow ϕ: $\mathbf{R}^2 \times \mathbf{R} \to \mathbf{R}^2$ which sends (x, t) to the point $g(t)$ where g is that solution of the system for which $g(0) = x$. We associate with this flow the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and ψ denote the set of such flows by Φ. For pictures of these flows see, for example, [**1**]. From now on, all flows considered are taken to be in Φ.

An *orbit* of a flow ϕ is a set $\{\phi(x,t) : t \in \mathbf{R}\}$ for some x. By a (ϕ, ψ)-*mapping*, where ϕ and ψ are flows, we mean a mapping of the plane which sends each orbit of ϕ onto an orbit of ψ. Flows ϕ and ψ are said to be *orbit-equivalent* if there is a (ϕ, ψ)-homeomorphism of the plane onto itself.

We give definitions of six types of flow which together make up the whole of Φ, and we subsequently show that the six types are the equivalence classes under orbit-equivalence. A non-singular matrix gives a *centre* if it has purely imaginary eigenvalues, a *saddle* if the determinant is negative, and a *topological source* otherwise. The flow given by a matrix of rank 1 has a line of equilibrium points, the other orbits forming a set of parallel lines; if these lines are parallel to the line of equilibrium points we call the flow a *shear*, otherwise a *line*. It is easy to show that, in the latter case only, the matrix has a non-zero eigenvalue. The *zero* flow, given by the zero matrix, is the only member of the sixth type. Consequently, in terms of the coefficients of the matrix, the type of a flow is determined by the sign, positive, negative or zero, of $ad - bc$, and whether or not $a + d$, $b - c$ are zero.

Flows in different types cannot be orbit-equivalent because they provide different sets of answers to the following three questions: Are there infinitely many point or-

---

D. M. Jordan studied category theory under A. Fröhlick at King's College, London, and he currently is a Lecturer at Hull.

H. L. Porteous recently finished his Ph. D. studies at the Univ. of Warwick under D. B. A. Epstein and M. Shub. He was a Fullbright Scholar at Berkeley, and has held positions at Hull and Liverpool Universities. *Editor.*

bits? Is there an orbit which is homeomorphic to $\mathbf{R}$? Is there an orbit which is homeomorphic to $\mathbf{R}$ and is a closed subset of the plane? It is more convenient to postpone the proof that flows of the same type are orbit-equivalent.

Orbit-equivalence gives a rather coarse classification, and a topological source is further classified by the number of its straight orbits: a *spiral* has no straight orbits, a *node* has four, a *degenerate node* has two and a *star* has infinitely many. Because of the correspondence between the straight orbits of a flow and the eigenvectors of the associated matrix, we can determine to which of these new types a topological source belongs if we know the sign of $(a+d)^2-4(ad-bc)$ and whether or not $b-c$ is zero.

We search for a relation whose equivalence classes are the nine types. A fairly subtle test for such a relation is that it distinguishes between stars and spirals and yet identifies all spirals. We show below that stars are distinguished from spirals by the relation which holds between flows ϕ and ψ when there is a (ϕ,ψ)-diffeomorphism. Nevertheless this relation fails: to see this we introduce the useful relation of *orbit-similarity*. Flows ϕ and ψ are *orbit-similar* if there is a (ϕ,ψ)-homeomorphism which is linear. Accordingly, two matrices give orbit-similar flows if and only if there is a non-zero multiple of one which is similar to the other. Now suppose that $\phi,\psi \in \Phi$ and that f is a (ϕ,ψ)-diffeomorphism. Then it turns out that ϕ and ψ are orbit-similar, the differential of f at the origin being a linear (ϕ,ψ)-homeomorphism: we have failed to find a very simple proof of this plausible assertion but give a proof in section 4. In consequence, although stars are distinguished from spirals, a classification based on diffeomorphisms has infinitely many classes of nodes, saddles and spirals.

We call a mapping *dependable* if it preserves linear dependence; in the case of a plane homeomorphism this means that lines through the origin are sent into lines through the origin. For any flow ϕ in Φ and any t in $\mathbf{R}$, the mapping $x \rightsquigarrow \phi(x,t)$, being a linear homeomorphism, is dependable.

We define flows ϕ,ψ to be of the same *orbit-type* if there is a dependable (ϕ,ψ)-homeomorphism. Thus flows in Φ of the same orbit-type are orbit-equivalent and have the same number of straight orbits. Because linear homeomorphisms are dependable, orbit-similar flows are of the same orbit-type. Inspection of the similarity classes of the 2×2 real matrices shows that any two centers are orbit-similar, as are any two shears, lines, degenerate nodes or stars. Consequently we establish that there are exactly nine orbit-types with the following proofs that the nodes, spirals, and saddles, each form a single orbit-type.

Again, inspection of similarity classes shows that each node is orbit-similar to the flow given by a diagonal matrix which satisfies $d-a=1$ and $a>0$. Any two such flows ϕ,ψ have restrictions ϕ_S,ψ_S which are homeomorphisms from $S^1\times\mathbf{R}$ onto $\mathbf{R}^2\setminus\{0\}$, where S^1 is the unit circle. Because $a>0$ for both ϕ and ψ, mapping the origin to itself extends $\psi_S\phi_S^{-1}$ to a homeomorphism of the plane. Lines through the origin are fixed lines of this homeomorphism since both ϕ and ψ have the property

that, in polar coordinates, $\dot{\theta} = \frac{1}{2}\sin 2\theta$. By the construction of ϕ_S, ψ_S, we see that $\psi_S\phi_S^{-1}$ maps orbits of ϕ onto orbits of ψ. The same argument applies to the spirals because each spiral is orbit-similar to a flow for which $\dot{\theta} = 1$ and $a > 0$.

Further inspection shows that any two saddles are orbit-similar to flows ϕ and ψ given by

$$\begin{pmatrix} 1 & 0 \\ 0 & -\lambda \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & -\mu \end{pmatrix}$$

respectively, where $\lambda, \mu > 0$. Put

$$W_1 = \{(x_1, x_2) : x_1 \geqq x_2 > 0\}, \qquad W_2 = \{(x_1, x_2) : x_2 \geqq x_1 > 0\},$$

denote the closures of W_1 and W_2 by C_1 and C_2, and let L be the line common to W_1 and W_2. The flows ϕ and ψ have restrictions ϕ_L, ψ_L which are homeomorphisms from $L \times \{t : t \geqq 0\}$ onto W_1. Since for each t in $\mathbf{R}$ the mappings $x \rightsquigarrow \phi(x,t)$ and $x \rightsquigarrow \psi(x,t)$ are dependable homeomorphisms, it follows that $\psi_L\phi_L^{-1}$ is dependable. Also $\dot{x}_1 = x_1$ for both ϕ and ψ ; so $\psi_L\phi_L^{-1}$ keeps vertical lines fixed. Consequently the identity on the x_1-axis extends $\psi_L\phi_L^{-1}$ to a dependable homeomorphism of C_1.

On C_2 we can use a similar construction since the orbits of ϕ and ψ are respectively those of

$$\begin{pmatrix} -1/\lambda & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} -1/\mu & 0 \\ 0 & 1 \end{pmatrix}.$$

These homeomorphisms on C_1 and C_2, both being the identity on L, produce a dependable homeomorphism of the first quadrant, which maps an orbit of ϕ meeting L at x to the orbit of ψ through x. The required plane homeomorphism is now obtained by symmetry.

With our results on orbit-types we complete the proof that there are just six equivalence classes under orbit-equivalence by showing that any two topological sources are orbit-equivalent. Matrix inspection shows that two such flows are orbit-similar to flows ϕ and ψ, whose orbits, apart from the equilibrium orbit, go outwards from the origin and meet S^1 just once. As before, ϕ and ψ have restrictions ϕ_S and ψ_S such that mapping the origin to itself extends $\psi_S\phi_S^{-1}$ to a (ϕ,ψ)-homeomorphism.

We now distinguish between inward and outward flows. Each orbit of a flow ϕ is made into a directed set by the relation, which we call the *direction*, in which $\phi(x,t) \geqq x$ whenever $t \geqq 0$. The direction is therefore an order on all the orbits, apart from the exceptional, periodic orbits. Flows ϕ and ψ are of the same *directed-orbit-type* if there is a dependable (ϕ,ψ)-homeomorphism which is direction preserving. Thus a topological source may be a *stable* or *unstable* node, degenerate node or spiral, or a *source* or a *sink*, and a line may be stable or unstable. Now if flows ϕ and ψ arise from matrices kA and PAP^{-1}, then

$$P\phi(x,t) = \psi(Px, kt),$$

showing that ϕ and ψ are of the same directed-orbit-type if $k > 0$. With our previous discussion this shows that the shears, saddles and centres form single directed-orbit-types, making just fourteen types in all. Because any eigenvalue of the matrix of a topological source has the same sign as $a + d$, it follows that $a + d < 0$ for a sink or stable flow and $a + d > 0$ for a source or unstable flow.

In the maps of Φ described below, six of the directed-orbit-types split further, each into a clockwise and an anti-clockwise component. It is straightforward to show that the resulting twenty types are obtained by defining ϕ and ψ to be of the same type if there is a direction preserving dependable (ϕ,ψ)-homeomorphism which preserves the sign of $\dot{\theta}(x)$ for all non-zero x. Because $\dot{\theta}(x)$ has the same sign as $cx_1^2 + (d - a)x_1x_2 - bx_2^2$ when $x \neq 0$, it follows that $b - c > 0$ for a clockwise flow and $b - c < 0$ for an anticlockwise flow.

An alternative aproach for both the preceding concepts is to use orientation, first for the orbits, then for the plane; the same classification results. We feel, however, that a precise formulation of this would be unnecessarily elaborate here.

2. Topologies. Up to now we have regarded Φ just as a set; by giving it additional structure, we can describe in more detail how it splits into directed-orbit-types. In fact we put a topology on it. There are several reasonable ways of doing this, but it turns out, because of the special properties of the flows in Φ, that the resulting topologies are all the same; we devote the rest of this section to showing this.

We call (a,b,c,d) the *coefficient vector* of the flow arising from the system

$$\dot{x}_1 = ax_1 + bx_2, \qquad \dot{x}_2 = cx_1 + dx_2\,.$$

We give Φ the *coefficient topology* by requiring that the correspondence between a flow and its coefficient vector is a homeomorphism. The distance $d(\phi,\psi)$ between flows ϕ and ψ is defined to be the distance between their coefficient vectors.

Next, define a subset N of Φ to be a neighbourhood of a flow ϕ if for some ε, R, T, where $\varepsilon > 0$, N contains every flow ψ such that

$$\left|\psi(x,t) - \phi(x,t)\right| < \varepsilon,$$

whenever $|x| \leqq R$ and $|t| \leqq T$. This yields the compact open topology on Φ regarded as a set of functions from $\mathbf{R}^2 \times \mathbf{R}$ to $\mathbf{R}^2$. Now the mapping $(\phi,(x,t)) \rightsquigarrow \phi(x,t))$ from $\Phi \times (\mathbf{R}^2 \times \mathbf{R})$ to $\mathbf{R}^2$ is continuous in the coefficient topology: this is a particular case of a general result frequently referred to as continuity of the solution with initial conditions and parameters. We deduce that the coefficient topology is larger than the compact open topology either from [**2**] or by using the uniform continuity of the mapping$(\phi,(x,t)) \rightsquigarrow \phi(x,t)$ on

$$\{\psi : d(\psi,\phi) \leqq 1\} \times \{(x,t) : |x| \leqq R, \quad |t| \leqq T\} \quad \text{for any} \quad R, T.$$

We now show that the coefficient topology is smaller than the compact open topology. For any flow ϕ write $\dot{\phi}(x,t)$ for the derivative at t of the mapping $t \rightsquigarrow \phi(x,t)$. Thus, if ϕ has matrix A, then $\dot{\phi}(x,t) = A\phi(x,t)$ for all x and t. Suppose now that ϕ and ψ are flows which satisfy $d(\psi,\phi) \geqq 2\eta > 0$, let A and B be the matrices of ϕ and ψ, and let E equal $B - A$. It follows that $\|E\| > \eta$. Let x be a point such that $|x| = 1$ and $|Ex| = \|E\|$, and take a neighbourhood $N(x, 2\delta)$ of x where $2\delta < \frac{1}{3}$ and $2\delta\|A\| < \frac{1}{6}\eta$. Accordingly, if $w \in N(x,2\delta)$, then

$$|Ew - Ex| < \tfrac{1}{3}\|E\|$$

and

$$|Aw - Ax| < \tfrac{1}{6}\eta.$$

Hence, if $w, z \in N(x, 2\delta)$, then

$$\begin{aligned} |(Bw - Az) - Ex| &\leqq |Aw - Az| + |Ew - Ex| \\ &< \tfrac{1}{3}\eta + \tfrac{1}{3}\|E\| \\ &< \tfrac{2}{3}\|E\|. \end{aligned}$$

Thus, whenever $\psi(y,t)$, $\phi(x,t) \in N(x,2\delta)$, we have

$$|(\dot{\psi}(y,t) - \dot{\phi}(x,t)) - Ex| < \tfrac{2}{3}\|E\|.$$

Choosing T so that $T > 0$ and $\exp(\|A\|T) - 1 < \delta$ ensures that $\phi(x,t) \in N(x,\delta)$ whenever $|t| \leqq T$. From the previous inequality we therefore deduce that either

$$|\psi(x,T) - \phi(x,T)| > \tfrac{1}{3}T\|E\|,$$

or

$$|\psi(x,t) - \phi(x,t)| > \delta$$

for some t satisfying $|t| \leqq T$. Thus any neighbourhood in the coefficient topology is a neighbourhood in the compact open topology, and the two topologies are the same.

From [2] again it now follows that another way of characterizing the coefficient topology on Φ is as the smallest topology which makes the mapping $(\phi,(x,t)) \rightsquigarrow \phi(x,t)$ continuous.

Alternatively, let a subset N of Φ be a neighbourhood of ϕ if for some ε, R, T, where $\varepsilon > 0$, N contains every flow ψ such that

$$|\psi(x,t) - \phi(x,t)| + |\dot{\psi}(x,t) - \dot{\phi}(x,t)| < \varepsilon,$$

whenever $|x| \leqq R$ and $|t| \leqq T$. It is clear that every neighbourhood in the coefficient topology is a neighbourhood in this new topology. Since, in the coefficient topology, the mapping $(\phi,(x,t)) \rightsquigarrow \dot{\phi}(x,t)$ is continuous, it is uniformly continuous on

$$\{\psi : d(\psi,\phi) \leqq 1\} \times \{(x,t) : |x| \leqq R, \quad |t| \leqq T\}$$

for any R, T; consequently the coefficient topology is larger than, and so the same as, the new topology.

It is not satisfactory to give Φ the smallest topology making $\phi \rightsquigarrow \phi(x, t)$ continuous for each (x, t). This, the pointwise topology on Φ, fails to make the mapping $(\phi, (x, t)) \rightsquigarrow \phi(x, t)$ continuous: to see this we show that the pointwise topology on Φ is strictly smaller than the coefficient topology. Let D be the set of pairs (X, ε) where X is a finite subset of $\mathbf{R}^2 \times \mathbf{R}$ and $\varepsilon > 0$. D is directed by saying that $(X, \varepsilon) \leqq (Y, \delta)$ if and only if $X \subseteq Y$ and $\varepsilon \geqq \delta$. It is easy to show that for each member (X, ε) of D there is a centre whose distance from the zero flow is at least 1 such that $|\phi(x, t) - x| < \varepsilon$ for all (x, t) in X. It follows that this net converges to the zero flow in the pointwise topology but not in the coefficient topology .

3. Maps. Flows have the same directed orbits if their coefficient vectors are on the same open ray from the origin in $\mathbf{R}^4$. Since each such ray meets S^3, the unit sphere in $\mathbf{R}^4$, just once, we see how Φ splits into directed-orbit-types if we describe the splitting of the set Ψ of flows whose coefficient vectors are in S^3. We give first a two-dimensional map of Ψ and later a more informative three-dimensional map.

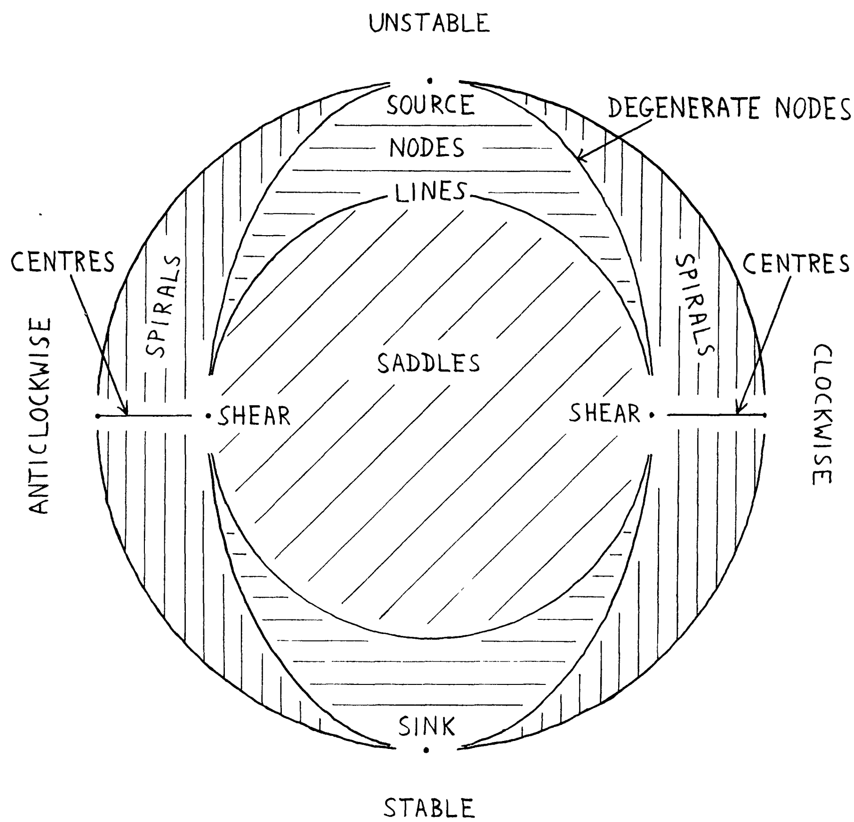


FIG. 1

Let $v(\phi)$ be the coefficient vector of a flow ϕ, let k be the orthogonal transformation

$$(a,b,c,d) \rightsquigarrow (b-c, a+d, b+c, a-d)/\sqrt{2}$$

and let p be the orthogonal projection $(u_1,u_2,u_3,u_4) \rightsquigarrow (u_1,u_2)$. Define the continuous mapping α from Ψ to the closed unit plane disc D^2 by taking $\alpha(\phi)$ to be $p(k(v(\phi)))$. We have shown that, if ϕ has coefficient vector (a,b,c,d), then the directed-orbit-type of ϕ is determined by the signs of $ad-bc$, $(a+d)^2-4(ad-bc)$, $b-c$ and $a+d$. Thus, if $\phi \in \Psi$ and $k(v(\phi)) = (u_1,u_2,u_3,u_4)$, the directed-orbit-type of ϕ is determined by the signs of $2u_1^2+2u_2^2-1$, $2u_1^2+u_2^2-1$, u_1 and u_2. Figure 1 shows the directed-orbit-type of the flows mapped by α to each point of D^2. The following properties of α can be readily checked. Flows ϕ and ψ in Ψ are mapped to the same point of D^2 if and only if some rotation sends ϕ into ψ: that is, there is a plane rotation r about the origin such that $\psi(r(x),t) = r(\phi(x,t))$ for all x and t. A flow ϕ is mapped to the boundary of D^2 if and only if ϕ is isotropic: that is, every rotation about the origin sends ϕ into itself. Because all stars are isotropic, this two-dimensional map has the misleading feature that shears and stars are both represented by two points although their dimensions in Ψ are different.

The homeomorphism β from Ψ to $\mathbf{R}^3 \cup \{\infty\}$ is constructed by taking $\beta(\phi)$ to be the stereographic projection of $k(v(\phi))$ from $(0,0,0,1)$ to the hyperplane $u_4 = 0$, Figure 2 shows the directed-orbit-type of the flow mapped to each point of $\mathbf{R}^3$. The following properties of β can be readily checked. The matrix of ϕ has zero determinant if and only if $\beta(\phi)$ is on the torus obtained by rotating the vertical circle with centre $(\sqrt{2},0,0)$ and radius 1 about the u_3-axis. This torus together with its inside is the image of the whole of Ψ except the saddles. As before, isotropic flows are mapped to S^1, the unit circle in the u_1,u_2 plane. Let ϕ be a non-isotropic flow. Then the set of flows obtained by rotating ϕ is mapped to a circle which is linked with S^1 and lies in a plane through the u_3-axis; the u_3-axis together with ∞ is to be regarded as such a circle. The surface $2u_1^2+u_2^2-1=0$, which corresponds to matrices with a single eigenvalue, is projected stereographically into a surface which meets the plane $u_3=0$ in two circles of radius $\sqrt{2}$ with centres $(1,0,0)$ and $(-1,0,0)$.

Because of the close connection between the type of a flow and standard properties of the associated matrix, it is possible to regard both of the above maps as providing information about the linear plane endomorphisms in their own right.

4. The failure of diffeomorphisms. We now give the promised proof that if f is a (ϕ,ψ)-diffeomorphism, where $\phi,\psi \in \Phi$, then the differential of f at the origin is a (ϕ,ψ)-mapping. In fact suppose that ϕ and ψ arise from matrices A and B and let f be a (ϕ,ψ)-homeomorphism which is differentiable at the origin with a non-singular differential there. We will prove that this differential is a (ϕ,ψ)-mapping. It is straightforward to deduce this result from the special case, proved below, in which this differential is the identity mapping.

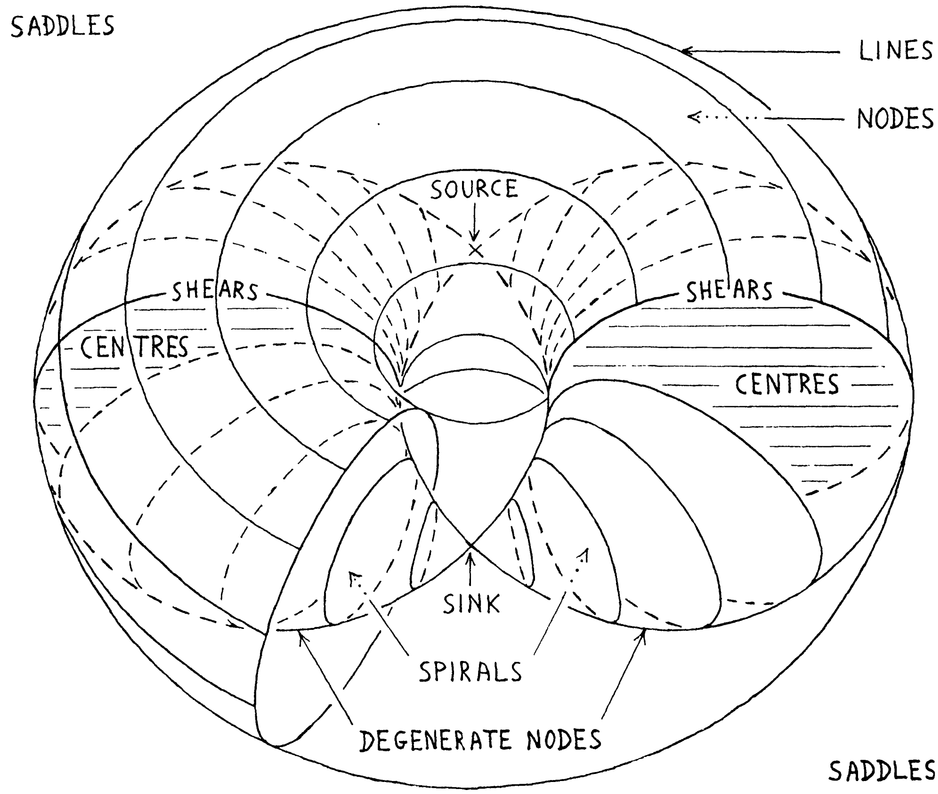


FIG. 2

We must prove that ϕ and ψ have the same orbits. Since the inverse of f is a (ψ, ϕ)-homeomorphism which has the identity differential at the origin, it is sufficient to prove that each non-equilibrium orbit of ϕ is contained in an orbit of ψ. Further, it is enough to prove that, if $Ax \neq 0$, then $\phi(x, t)$ is on the orbit of ψ through x for all sufficiently small t.

Suppose that $Ax \neq 0$. Because the inverse of f has the identity differential at the origin and maps equilibrium points of ψ to equilibrium points of ϕ, we deduce that $Bx \neq 0$. Hence there is a strictly positive number δ such that, if $w \in N(x, 2\delta)$, then

$$\left|Bw - Bx\right| \leqq \tfrac{1}{2}\left|Bx\right|.$$

Because the mapping $t \rightsquigarrow \phi(x, t)$ is continuous, there is a number T such that $T > 0$ and, if $\left|t\right| \leqq T$, then $\phi(x, t) \in N(x, \delta)$. Take any ε which satisfies $0 < \varepsilon < \delta/(\left|x\right| + \delta)$. Then there is a neighbourhood N of the origin such that

$$\left|f(w) - w\right| \leqq \varepsilon\left|w\right|$$

or all w in N. Choose λ so that $N(\lambda x, \lambda\delta) \subseteq N$. Because $\phi(\lambda x, t) = \lambda\phi(x, t)$, it follows that if $|t| \leqq T$, then

$$\begin{aligned} |f\phi(\lambda x, t) - \phi(\lambda x, t)| &\leqq \varepsilon|\phi(\lambda x, t)| \\ &< \varepsilon\lambda(|x| + \delta) < \lambda\delta \end{aligned}$$

and

$$|\phi(\lambda x, t) - \lambda x| < \lambda\delta.$$

Hence $f\phi(\lambda x, t) \in N(\lambda x, 2\lambda\delta)$ whenever $|t| \leqq T$, and our first inequality gives

$$|Bf\phi(\lambda x, t) - B\lambda x| \leqq \tfrac{1}{2}|B\lambda x|$$

whenever $|t| \leqq T$. Because f is a (ϕ, ψ)-homeomorphism, there is a homeomorphism h of the real line such that

$$f\phi(\lambda x, t) = \psi(f(\lambda x), h(t))$$

for all t. Hence, writing $f(\lambda x) = y$, the previous inequality gives

$$|\psi(y, h(t)) - B\lambda x| \leqq \tfrac{1}{2}|B\lambda x|$$

whenever $|t| \leqq T$. We deduce that

$$|\psi(y, h(t)) - y| \geqq \tfrac{1}{2}|h(t)|\,|B\lambda x|$$

whenever $|t| \leqq T$, whence

$$\begin{aligned} |h(t)| &\leqq 2|\psi(y, h(t)) - y|/|B\lambda x| \\ &\leqq 2(|\psi(y, h(t)) - \lambda x| + |y - \lambda x|)/|B\lambda x| \\ &< 8\lambda\delta/|B\lambda x| \\ &= 8\delta/|Bx|. \end{aligned}$$

Since ψ is in Φ, the mapping $w \rightsquigarrow \psi(w, s)$ is linear, with norm m say; moreover, m is a continuous function of s. Let M be $\sup\{m: |s| \leqq 8\delta/|Bx|\}$, and t be any number such that $|t| \leqq T$. Then

$$\begin{aligned} |\psi(y, h(t)) - \psi(\lambda x, h(t))| &\leqq M|y - \lambda x| \\ &\leqq M\varepsilon|\lambda x|. \end{aligned}$$

Hence

$$\begin{aligned} |\phi(\lambda x, t) - \psi(\lambda x, h(t))| &\leqq |\phi(\lambda x, t) - f\phi(\lambda x, t)| + |\psi(y, h(t)) - \psi(\lambda x, h(t))| \\ &< \varepsilon\lambda(|x| + \delta) + M\varepsilon\lambda|x|, \end{aligned}$$

and so

$$|\phi(x, t) - \psi(x, h(t))| < \varepsilon(|x| + \delta + M|x|).$$

Thus $\phi(x,t)$ is on the orbit of ψ through x whenever $|t| \leqq T$.

References

1. L. S. Pontryagin, Ordinary differential equations, Pergamon, Long Island City, N. Y., 1962.
2. J. L. Kelley, General topology, Van Nostrand, Princeton, N. J., 1955, 221–225.

RECONSTRUCTING THE HISTORY AND GEOGRAPHY OF AN EVOLUTIONARY TREE

DAVID SANKOFF, Université de Montréal

1. Introduction. In the process of phylogenesis a species splits into two or more populations which evolve independently into distinct varieties. Later, any of these may in turn split. As time progresses, current populations which stem from different branches of an earlier split may constitute distinct species, genera, families, etc. Biologists have traditionally represented this process in terms of tree diagrams, as in Figure 1a. At each time $t \in [-T, 0]$ where $-T$ is the date of the first split, and the present is time zero, a tree consists of a number of populations, each of which is the forerunner or ancestor of a certain subset of the present-day populations (e.g., Figure 1b).

DEFINITION 1. An evolutionary tree on a finite set S is a family $\{\mathscr{P}_t\}_{-T}^{0}$ of partitions of S, where

$$\mathscr{P}_{-T} = \{S\},\ \mathscr{P}_0 = \{\{X\} \mid X \in S\},$$

$$-T \leqq t \leqq u \leqq 0 \Leftrightarrow \mathscr{P}_u \text{ is a refinement of } \mathscr{P}_t$$

and $\lim_{t \uparrow u} \mathscr{P}_t = \mathscr{P}_u$.

DEFINITION 2. Let $\{\mathscr{P}_t\}_{-T}^{0}$ be an evolutionary tree on S. Every subset $X \subseteq S$ where $X \in \mathscr{P}_t$ for some $t \in [-T, 0]$, denotes a **population** in the tree. We shall have occasion to distinguish X_t, population X at time t, from X_u, the same population at time u, for $X \in \mathscr{P}_t \cap \mathscr{P}_u$. If $t < u$, we say X_t is **ancestral** to X_u. A population X is ancestral to a population Y if $Y \subset X$, and then we may also say X_t is ancestral to Y_u for all $\mathscr{P}_t$, $\mathscr{P}_u$ where $X \in \mathscr{P}_t$, $Y \in \mathscr{P}_u$.

The major problem in genetic taxonomy is as follows. Given a set S of genetically

David Sankoff received his McGill Univ. Ph.D. in 1969 under D. A. Dawson. His unusually wide background in statistics, biology, social sciences, and linguistics includes research assistantships in mathematics, sociology, anthropology, and anatomy at McGill and field work over several years in New Guinea. He is a member of the Univ. de Montréal Centre de recherches mathématiques. *Editor.*

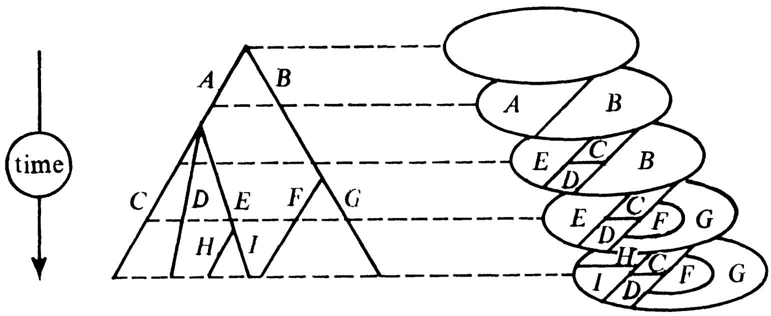


FIG. 1a FIG. 1b

related, currently existing (at time $t = 0$), populations, how can their evolutionary tree be deduced? In the next section we study a model of genetic divergence where, once split apart, populations evolve completely independently of one another. In this case reconstruction of the evolutionary tree from data on the existing populations is quite easy. This model is appropriate for trees which contain different genera, families, classes, etc., which *do* evolve relatively independently.

For evolutionary trees of populations which all belong to the same *species*, however, the problem is much more difficult since there may be interactions, i.e., interbreeding, between the various branches. In Section III we develop a model for this more interesting genetic divergence process, in terms of which we can solve the reconstruction problem.

2. Genetic divergence; independent populations. The similarity between two populations,

$$s(X_t, Y_u) = s(Y_u, X_t) \geqq 0,$$

is measured by the proportion of gene **types** they have in common. More specifically, there is some fixed set Γ of genetic **sites**, and at each site the two populations either have the same gene type or two completely different types. (We ignore the small proportion of gene sites for which there may be different types within a single population.) We assume Γ sufficiently large that we can neglect statistical fluctuation in the dynamic models we shall discuss.

Note that

$$s(X_t, X_t) = 1. \tag{1}$$

The simplest quantitative model of evolutionary divergence posits that in Γ, each site has a constant probability r per unit time of undergoing a replacement event. Then the probability of a type remaining unreplaced over a time interval of length $u - t$ satisfies the differential equation

$$\frac{d\Pr(u-t)}{du} = -r\Pr(u-t), \tag{2}$$

(see Feller, [1] Chapter XVII). The assumption that Γ is large may be rephrased mathematically as an assumption that the proportion of sites escaping replacement will also satisfy this equation. (Were Γ small, (2) would hold only for the expected value of the proportion.) Under the hypothesis that once replaced, a type can never recur, it follows that similarity also obeys (2). In other words, for X ancestral to Y (including the case when $X = Y$ but $u \geqq t$),

$$\frac{ds(X_t, Y_u)}{du} = -rs(X_t, Y_u), \tag{3}$$

from which we immediately derive, for initial condition (1):

PROPOSITION 1. *For X ancestral to Y, $s(X_t, Y_u) = \exp[-r(u-t)]$.*

Under the further hypothesis that a new type cannot occur as an innovation in two or more populations, and interpreting independent evolution in terms of probabilistic independence, we have the following more general statement:

PROPOSITION 2. *For all X and Y*

$$s(X_t, Y_u) = \exp[-r(v-t)]\exp[-r(v-u)],$$

where v is the latest point of time at which there exists a population ancestral to both X_t and Y_u.

Proof. For X_t ancestral to Y_u, it is clear that $v = t$, in which case we use Proposition 1. Likewise for Y_u ancestral to X_t.

In all other cases there will be a most recent population Z ancestral to both X and Y. Let

$$v = \max\{\tau \mid Z \in \mathscr{P}_\tau\}.$$

The maximum exists because of the limit assumption in Definition 1. Then

$$s(Z_v, X_t) = \exp[-r(v-t)],$$
$$s(Z_v, Y_u) = \exp[-r(v-u)].$$

By independence, the probability of a site being unaffected by replacement both between Z_v and X_t, and between Z_v and Y_u, is the product of the probabilities for the individual events. The same product relation holds for proportions of types unreplaced, by our assumption about Γ. The hypothesis of uniqueness of innovation ensures that the coefficient of similarity between X_t and Y_u will be precisely the proportion of sites unaffected by replacement in both evolutionary branches. Hence

$$s(X_t, Y_u) = s(Z_v, X_t)s(Z_v, Y_u)$$

which proves the proposition.

An **ultrametric** space $(\boldsymbol{S}, d)$ is a metric space where, for W, X, $Y \in \boldsymbol{S}$,

$$d(X, Y) \leqq \max\{d(X, W), d(Y, W)\}. \tag{4}$$

At time zero, i.e., the present, let $\boldsymbol{S}$ be the set of populations currently representative of a given evolutionary tree. Without ambiguity, we can write X for $\boldsymbol{X}_0 = \{X\}$.

For X, $Y \in \boldsymbol{S}$ let $v(X, Y)$ be the time of the most recent common ancestor of X and Y as defined in Proposition 2.

PROPOSITION 3. *The pair* $(\boldsymbol{S}, -v)$ *is an ultrametric space.*

Proof. Clearly $-v(X, Y) = 0$ if and only if $X = Y$; and $-v(X, Y) = -v(Y, X)$. It remains to prove (4), the ultrametric inequality (which implies the triangle inequality required of a metric). Suppose it does not hold and for some W, X, $Y \in \boldsymbol{S}$

$$-v(X, Y) > \max\{-v(X, W), -v(Y, W)\}.$$

Then X and W have a more recent common ancestor population $\boldsymbol{Z}^{(1)}$ than do X and Y, and Y and W have a more recent common ancestor $\boldsymbol{Z}^{(2)}$ than do X and Y. But by Definitions 1 and 2, the ancestors of W form a nested sequence of subsets of $\boldsymbol{S}$. Therefore one of $\boldsymbol{Z}^{(1)}$ or $\boldsymbol{Z}^{(2)}$ must be a common ancestor to both X and Y, contrary to our supposition. Hence the ultrametric inequality holds.

PROPOSITION 4. *Let* $\boldsymbol{S}$ *represent a finite set of populations existing at time zero. An ultrametric* d *on* $\boldsymbol{S}$ *determines a unique evolutionary tree where, if* X, $Y \in \boldsymbol{S}$, *then* $-d(X, Y)$ *is the date of the most recent population ancestral to both* X *and* Y.

Proof. There are a finite number of different values of d, say $0 < d_1 < \cdots < d_m = -T$. For each $X \in S$, consider the nested sequence of sets

$$\boldsymbol{X}_0 = \{X\} \subseteq \cdots \subseteq \boldsymbol{X}_{-d_i} = \{Y \in \boldsymbol{S} \mid d(X, Y) \leqq d_i\} \subseteq \cdots \subseteq \boldsymbol{X}_{-T} = \boldsymbol{S}.$$

The ultrametric inequality (4) assures, for any two such sequences $\boldsymbol{X}_0, \cdots, \boldsymbol{X}_{-T}$ and $\boldsymbol{Y}_0, \cdots, \boldsymbol{Y}_{-T}$, there is an integer p satisfying

$$\begin{aligned} \boldsymbol{X}_t \cap \boldsymbol{Y}_t &= \varnothing \quad \text{for } t = 0, -d_1, \cdots, -d_p \\ \boldsymbol{X}_t = \boldsymbol{Y}_t & \qquad\quad \text{for } t = -d_{p+1}, \cdots, -d_m. \end{aligned} \tag{5}$$

Let $\mathscr{P}_0 = \{\{X\} \; X \in \boldsymbol{S}\}$, and, for $t = d_i$, let $\mathscr{P}_t$ be the set of distinct $\boldsymbol{X}_t$. For $-d_i < t \leqq -d_{i-1} = u$, let $\mathscr{P}_t = \mathscr{P}_u$, for $i = 1, \cdots, m$. From (5) it follows that $\{\mathscr{P}_t\}_{-T}^{0}$ satisfies Definition 1. For any X, $Y \in \boldsymbol{S}$, our construction assures that X and Y are in the same element of $\mathscr{P}_t$ for t up to and including $t = -d(X, Y)$. This is precisely the ancestry condition required by the theorem, and it uniquely determines $\{\mathscr{P}_t\}_{-T}^{0}$.

Propositions 2–4 provide a solution to the reconstruction problem. The biologist first measures the similarities between the populations in $\boldsymbol{S}$. Using the special case of Proposition 2 where $u = t = 0$, he solves

$$v(X, Y) = \frac{1}{2r} \log s(X, Y), \quad \text{for all } X, Y \in \boldsymbol{S}.$$

By Proposition 3, $(\boldsymbol{S}, -v)$ is an ultrametric space which, by Proposition 4, uniquely determines the evolutionary tree of $\boldsymbol{S}$. In fact, the proof of this latter proposition includes a construction of the tree.

3. Divergence with interaction. To be able to treat the case where changes in one population can be influenced by another, we add a geographical dimension to our hitherto purely historical considerations. At any point $t \in [-T, 0]$, each population in $\mathscr{P}_t$ will be associated with a face of a planar graph $\mathscr{M}_t$. This is illustrated in Figure 2.

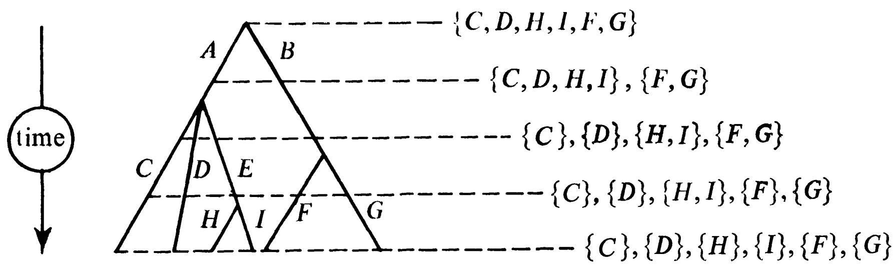


FIG. 2

For any tree $\mathscr{M}_{-T}$ is a loop, or degenerate graph consisting of one edge, one face and no vertices, as in Figure 2. If the split at time $-T$ is into two populations, then $\mathscr{M}_t$, for t immediately after $-T$, is a graph consisting of two faces, three edges (two exterior, one interior) and two vertices. Whenever a population splits into n fragments, the face corresponding to it is subdivided into two portions, then one of these two is further subdivided, then one of the resulting three is chosen for further subdivision, and so on, until an n-way fragmentation is achieved. The subdivision of a face is accomplished by choosing any two distinct edges bordering that face, placing a new vertex midway along each of these edges and joining the two vertices with a new edge. Alternatively, if the face has an exterior border (the ocean!), two new vertices on this single edge may be joined by a new edge.

DEFINITION 3. A geography associated with an evolutionary tree $\{\mathscr{P}_t\}_{-T}^{0}$, is a family of planar graphs $\{\mathscr{M}_t\}_{-T}^{0}$ where there is 1-1 correspondence between the populations of $\mathscr{P}_t$ and the faces of $\mathscr{M}_t$, satisfying

(a) $\mathscr{M}_{-T}$ is a loop,

(b) for two successive refinements $\mathscr{P}_t$, $\mathscr{P}_u$

$$\boldsymbol{X} \in \mathscr{P}_t,\ \boldsymbol{X}^{(1)}, \cdots, \boldsymbol{X}^{(n)} \in \mathscr{P}_u,\ \boldsymbol{X} = \bigcup_{i=1}^{n} \boldsymbol{X}^{(i)}$$

$\Rightarrow \mathscr{M}_u$ is derived from $\mathscr{M}_t$ by the subdivision of the face corresponding to $\boldsymbol{X}$ into the faces corresponding to $\boldsymbol{X}^{(1)}, \cdots, \boldsymbol{X}^{(n)}$.

This is the simplest way of constructing planar graphs by extension and hence is the simplest model of the evolution of territorial configurations of related populations.

How do populations interact? Instead of just replacing types at sites in Γ with completely new types, we now allow, in addition, the adoption of types from neighboring populations. Two neighboring populations are, of course, populations whose corresponding faces share an edge.

If X and Y are neighbors, we write

$$X \in N_Y \Leftrightarrow Y \in N_X.$$

We can construct models where the total replacement rate is constant but the proposition of adoptions depends on the number of neighbors, other models where new replacements occur at a constant rate but the adoption rate depends on the number of neighbors, or models where the adoption rate is constant. Mathematically speaking, these all lead to the same type of problem, and so we study just the last one.

We shall describe the genetic divergence process between two successive splits. In this interval $\mathscr{M}_t$ and $\mathscr{P}_t$ are fixed, so we can suppress the time subscripts on populations without risking ambiguity.

For each population X, we assume a probability rate r for new replacements as before, and probability rate $a/k(X)$ for adoptions from each of its $k(X)$ neighbors. Suppose $X \in N_Y$. Then $ds(X_t, Y_t)/dt$, the rate of change in the similarity between the two simultaneously evolving populations X and Y, has several components. There is the change due to new replacements, into X and into Y; the change due to adoptions from X into Y and vice-versa; and finally the change due to adoptions into X and Y from their other neighbors. For the first component, the same arguments which justify (2) and (3) in the case of a single evolutionary line, also imply that the change rate due to new replacements into the *two* populations is $-2rs(X_t, Y_t)$. (Were all other components zero, this could also be derived directly from Proposition 2, where $t = u$.) For the next component, the total adoption rate between X and Y is $a(1/k(X) + 1/k(Y))$ but a proportion $s(X_t, Y_t)$ of types adopted are already identical in the two populations so that the change rate due to this process will be $(1 - s(X_t, Y_t))\, a(1/k(X) + 1/k(Y))$. In addition we must take into account adoptions from the remaining $k(X) - 1$ neighbors of X and the remaining $k(Y) - 1$ neighbors of Y. Adoptions from neighbors of X change the similarity at a rate

$$\frac{a}{k(X)-1} \sum_{Z \in N_X - Y} [(1 - s(X_t, Y_t))s(Y_t, Z_t) - s(X_t, Y_t)(1 - s(X_t, Z_t))], \tag{6}$$

following the same line of reasoning, and adoptions from neighbors of Y have an analogous effect, but with Y and X interchanged in (6).

Collecting terms, we find that

$$\frac{ds(X_t, Y_t)}{dt} = -\beta(X, Y)s(X_t, Y_t) + \alpha(X, Y), \tag{7}$$

where

$$\alpha(X, Y) = a\left\{\frac{1}{k(X)} + \frac{1}{k(Y)}\right\} + \frac{a}{k(X)-1}\sum_{Z\in N_X - Y} s(Y_t, Z_t)$$
$$+ \frac{a}{k(Y)-1}\sum_{Z\in N_Y - X} s(X_t, Z_t),$$

$$\beta(X, Y) = 2r + \alpha(X, Y) + \frac{a}{k(X)-1}\sum_{Z\in N_X - Y}(1 - s(X_t, Z_t)) \tag{8}$$
$$+ \frac{a}{k(Y)-1}\sum_{Z\in N_Y - X}(1 - s(Y_t, Z_t)).$$

For two populations X and Y which are not neighbors, coefficients $\alpha(X, Y)$ and $\beta(X, Y)$ are as in (8) but without the term $a(1/k(X) + 1/k(Y))$.

PROPOSITION 5. *Let $\{\mathscr{P}_t\}_{-T}^{0}$ be an evolutionary tree with associated geography $\{\mathscr{M}_t\}_{-T}^{0}$. If genetic divergence proceeds according to* (7), *then $\mathscr{P}_0$, $s(X_0, Y_0)$ for all X_0, $Y_0 \in \mathscr{P}_0$, and $\mathscr{M}_0$ uniquely determine the tree and its geography.*

Proof. The graph $\mathscr{M}_0$ summarizes all neighboring relations between populations in $\mathscr{P}_0$. These relationships are fixed as far back as $\mathscr{P}_t$ remains unchanged. Therefore we can write down equation (7) explicitly, with initial conditions $s(X_0, Y_0)$. The system of first-order equations so obtained satisfies conditions for a unique solution and can be solved by successive approximation. We write the solution as s'.

Suppose the most recent population split was at time v, when populations W and Z were formed from population $\{W, Z\}$. Immediately after v, and any time $s(W_t, Z_t)$ is close to 1,

$$\frac{ds(W_t, Z_t)}{dt} < 0,$$

as can be seen in (6) or (8). Thus, $s(W_t, Z_t) < 1$ on $(v, 0]$. But, by condition (1) at v,

$$\lim_{t\downarrow v} s(W_t, Z_t) = 1.$$

Then time v, and the populations W and Z can be found as

$$v = \max\{\tau \mid \exists X, Y \in \mathscr{P}_0, s'(X_\tau, Y_\tau) = 1\},$$

and $s' = s$ on $(v, 0]$.

The graph $\mathscr{M}_v$ is then constructed by deleting the edge between the faces corresponding to W and Z. Note that by continuity

$$\lim_{t\downarrow v} s(X_t, W_t) = \lim_{t\downarrow v} s(X_t, Z_t) \text{ for all } X \in S,$$

since s measures proportions of types shared by two populations.

We now have $\mathscr{P}_v$, $s(X_v, Y_v)$ for all X, $Y \in \mathscr{P}_v$, and $\mathscr{M}_v$. We can then set up a new system of equations (7) with initial conditions $s(X_v, Y_v)$ and solve as before. The new solution s'' will be valid as far back as the second most recent split, and so on. The generalization to n-way splits, and the case where more than one population splits at the same instant, are obvious. We continue the solution procedure until we have deleted the last non-exterior line of the graph, which gives us $-T$ and $\mathscr{M}_{-T}$. This construction, for which each step is uniquely determined, proves the proposition.

This last proposition means that a biologist, equipped with similarity data as well as a knowledge of the geographical configuration of a number of currently existing related populations, can reconstruct the entire evolutionary tree of the populations, as well as the geographical configuration at all times in $[-T, 0]$.

4. Discussion. There are a number of practical problems associated with the theory of both Section II and Section III. One is that Γ is too small to ignore statistical fluctuation. Another is that r and a are not universal constants but may change somewhat from site to site in Γ and from population to population. The hypotheses about non-recurrence of innovation are not always justified. When these factors are taken into account, the reconstruction methods we have described must be bolstered by search algorithms and statistical estimation. Some useful references are: Dayhoff [2], Sokal and Sneath [3], Lerman [4], and Jardine and Sibson [5].

References

1. William Feller, An Introduction to Probability Theory and Its Applications, 2nd ed, Wiley, New York, 1957.

2. M. O. Dayhoff, Computer analysis of protein evolution, Scientific American, 221 (1969) 86–95.

3. Robert R. Sokal and Peter H. Sneath, Principles of Numerical Taxonomy, Freeman, San Francisco, 1963, (2nd ed. forthcoming).

4. I. C. Lerman, Les Bases de la Classification Automatique, Gauthier-Villars, Paris, 1970.

5. N. Jardine and R. Sibson, Mathematical Taxonomy, Wiley, New York, 1971.

LIPSCHITZIAN POINTS

E. M. BEESLEY, University of Nevada, Reno, A. P. MORSE, University of California, Berkeley, and D. C. PFAFF, University of Nevada, Reno

1. Introduction. Notations are explained in the next two paragraphs and the first paragraph of Section 2.

Throughout we understand: that R is the set of real finite numbers; that Rp is the set of real finite positive numbers; that ω is the set of nonnegative integers; with fractions in mind, that F is the set of rational numbers; that J is the open

unit interval; and, for $r \in \mathrm{F}$, that den r is the smallest positive integer q such that, for some integer p, $r = p/q$.

We shall assume that W and w are such functions on J that, for each $x \in \mathrm{J}$,

$$W(x) = \{r \in \mathrm{F} : 0 \leqq r \leqq x\}$$

and

$$w(x) = \sum_{r \in W(x)} (\operatorname{den} r)^{-3}.$$

On page 408 of [**2**] we find the following theorem:

FORT'S THEOREM. *If f is on* R *to* R *with dense discontinuities, then the points at which f is differentiable form a set of the first category.*

Boas, [**1**, pp. 126–7] has reaffirmed this theorem and Heuer [**3**] has modified Fort's proof to obtain the result which follows.

HEUER'S THEOREM. *If f is on* R *to* R *with dense discontinuities and if $0 < \alpha \leqq 1$, then the set of points at which f satisfies a Hölder [or Lipschitz] condition of order α is of the first category.*

We find all three proofs convincing although we do notice that the second inequality of line 6 of [**1**] is miscast and should be reversed.

Nevertheless a recent note [**5**] includes a claim (now withdrawn [**6**]) that Fort's Theorem is invalid and the function w above is a counterexample to it. The supporting argument hinged on the assertion below.

ASSERTION 1.1. *If x is an irrational number in* J, *then $w'(x) = 0$.*

This assertion is not correct, but scrutiny of it has led us to the results below.

Fort's Theorem is of course a consequence of Heuer's Theorem which, in turn, is an immediate consequence of Theorem 2.1 below. The scope of Theorem 2.1 almost forces a simple proof upon us. Also in Section 2 are Theorem 2.2 and Application 2.3 which suggest that discontinuity here plays a less relevant role than one might think. Minor changes in the proof of 2.1 yield a proof of the more general 2.4.

In Section 3, which is independent of Section 2, we use Theorem 3.1 specifically to counter 1.1 and, unexpectedly, to obtain a short proof of the theorem below.

KHINCHIN'S THEOREM [**4**, p. 69]. *If f is such a function to* Rp *such that*

$$\sum_{q=1}^{\infty} f(q) < \infty,$$

then for almost all $x \in \mathrm{R}$ there are not infinitely many $r \in \mathrm{F}$ for which

$$|x - r| < \frac{f(\operatorname{den} r)}{\operatorname{den} r}.$$

2. Lipschitzian Structure. We now look at things a bit differently. If ρ_1 metrizes S_1, ρ_2 metrizes S_2, and f is a function on S_1 to S_2, then we agree that

$$\operatorname{Lip}\rho_1\rho_2 f = \{x \in S_1 \colon \text{there are finite positive numbers } M \text{ and } \delta \text{ for which } \rho_2(f(x),f(y)) \leqq M\cdot\rho_1(x,y) \text{ whenever } y \text{ is such that } \rho_1(x,y) < \delta\}.$$

We begin the proof of Theorem 2.1 by indicating, in effect, that there is no loss in generality in assuming that ρ_2 is bounded.

THEOREM 2.1. *If ρ_1 metrizes S_1, ρ_2 metrizes S_2, and f is a function on S_1 to S_2, then* $\operatorname{Lip}\rho_1\rho_2 f$ *is a countable union of closed ρ_1 sets.*

Proof. We can and do so let ρ_3 metrize S_2 that

$$\rho_3(x,y) = \frac{\rho_2(x,y)}{1+\rho_2(x,y)},$$

whenever $x \in S_2$ and $y \in S_2$. We check that

$$\operatorname{Lip}\rho_1\rho_2 f = \operatorname{Lip}\rho_1\rho_3 f. \tag{1}$$

Next, for $\nu \in \omega$, we let $A_\nu = \{x \in S_1 \colon \rho_3(f(x),f(y)) \leqq \nu\cdot\rho_1(x,y) \text{ whenever } y \in S_1\}$. Since ρ_3 is bounded, we infer that $\operatorname{Lip}\rho_1\rho_3 f = \bigcup_{\nu\in\omega} A_\nu$ and then use (1) to learn that

$$\operatorname{Lip}\rho_1\rho_2 f = \bigcup_{\nu\in\omega} A_\nu. \tag{2}$$

Now we assume $\nu \in \omega$ and that x belongs to the closure ρ_1 of A_ν. From A_ν we select a sequence ξ for which

$$\lim_n \rho_1(\xi_n, x) = 0. \tag{3}$$

Since $\rho_3(f(\xi_n),f(x)) \leqq \nu\cdot\rho_1(\xi_n,x)$ whenever $n \in \omega$, we learn from (3) that

$$\lim_n \rho_3(f(\xi_n),f(x)) = 0. \tag{4}$$

For each $y \in S_1$ we have: $\rho_3(f(\xi_n),f(y)) \leqq \nu\cdot\rho_1(\xi_n,y)$ whenever $n \in \omega$; and, because of this, (3), and (4),

$$\rho_3(f(x),f(y)) \leqq \nu\cdot\rho_1(x,y).$$

We conclude that $x \in A_\nu$.

Because of the above paragraph we see that A_ν is closed ρ_1 whenever $\nu \in \omega$, and complete our proof with the help of (2).

An immediate consequence of Theorem 2.1 is the following theorem:

THEOREM 2.2. *If ρ_1 metrizes S_1, ρ_2 metrizes S_2, f is on S_1 to S_2, and the complement of* $\operatorname{Lip}\rho_1\rho_2 f$ *with respect to S_1 is dense ρ_1, then* $\operatorname{Lip}\rho_1\rho_2 f$ *is of the first category ρ_1.*

APPLICATION 2.3. *If g is such a function on* R *that*

$$g(x) = \left(\frac{x}{1+|x|}\right)^{1/3} \text{ whenever } x \in \mathrm{R},$$

r is a sequence whose range is the rationals, f is such a function on R *that*

$$f(x) = \sum_{n \in \omega} \frac{g(x - r_n)}{2^n} \text{ whenever } x \in \mathrm{R},$$

and ρ is the usual metric for R, *then*: *f is bounded, absolutely continuous, and increasing*; Lip$\rho\rho f$ *is a set of the first category whose complement with respect to* R *is of Lebesgue measure zero.*

By making minor changes in the proof of 2.1, we can easily check the more general theorem which follows.

THEOREM 2.4. *If ϕ is a continuous function on $\{t : 0 \leqq t < \infty\}$,*

$$\phi(0) = 0,$$

$$\phi(t) > 0 \text{ whenever } t \in \mathrm{Rp},$$

$$\varliminf_{t \to \infty} \phi(t) > 0,$$

ρ_1 metrizes S_1, ρ_2 metrizes S_2, f is a function on S_1 to S_2, and

$$L = \{x \in S_1 : \text{there are finite positive numbers } M \text{ and } \delta \text{ for which } \rho_2(f(x), f(y)) \leqq M \cdot \phi(\rho_1(x, y)) \text{ whenever } y \text{ is such that } \rho_1(x, y) < \delta\},$$

then L is a countable union of closed ρ_1 sets.

Theorem 2.4 remains valid if we replace '$>$' by '$\geqq$' therein. The proof we have in mind avoids the introduction of ρ_3 but strikes us as more intricate than our present proof of 2.1.

3. Upper Derivatives of Certain Jump Functions. Rather than merely refute 1.1, we focus instead on the more general theorem below.

THEOREM 3.1. *If f is such a function to* Rp *that $\sum_{q=1}^{\infty} f(q) < \infty$, g is such a function that*

$$g(x) = \sum_{r \in W(x)} \frac{f(\operatorname{den} r)}{\operatorname{den} r}, \text{ whenever } x \in \mathrm{J},$$

$A = \{x \in \mathrm{J}$: *there are infinitely many $r \in \mathrm{F}$ for which*

$$|x - r| < \frac{f(\operatorname{den} r)}{\operatorname{den} r}\Big\},$$

$B = \{x \in \mathrm{J}$: *for each* $\lambda \in \mathrm{Rp}$, *there are infinitely many* $r \in \mathrm{F}$ *for which*

$$\lambda \cdot | x - r | < \frac{f(\operatorname{den} r)}{\operatorname{den} r} \Big\},$$

then:

(1) *if* $x \in A$, *then* x *is irrational and* $\overline{\lim\limits_{t \to x}} \dfrac{g(t) - g(x)}{t - x} \geqq \dfrac{1}{2}$;

(2) A *is of Lebesgue measure zero*;

(3) *if* $x \in B$, *then* $x \in A$ *and* $\overline{\lim\limits_{t \to x}} \dfrac{g(t) - g(x)}{t - x} = \infty$.

Proof. Letting

$$M = \sum_{q=1}^{\infty} f(q)$$

and observing, for $x \in \mathrm{J}$, that

$$g(x) = \sum_{r \in W(x)} \frac{f(\operatorname{den} r)}{\operatorname{den} r} \leqq \sum_{q=1}^{\infty} \frac{q \cdot f(q)}{q} = M < \infty,$$

we notice that g is a bounded increasing pure jump function on J to Rp. Since such a function must have zero derivative almost everywhere in J [**1**, p. 130], we see that (2) is a consequence of (1).

We see that (1) and (3) are consequences of the statement below.

STATEMENT. *If* $x \in \mathrm{J}$, $\lambda \in \mathrm{Rp}$, *and there are infinitely many* $r \in \mathrm{F}$ *for which*

$$\lambda \cdot |x - r| < \frac{f(\operatorname{den} r)}{\operatorname{den} r},$$

then x *is irrational and*

$$\overline{\lim_{t \to x}} \frac{g(t) - g(x)}{t - x} \geqq \frac{\lambda}{2}.$$

Proof. From F we can and do select such a univalent sequence y that, for $n \in \omega$,

$$0 < \lambda \cdot |y_n - x| < \frac{f(\operatorname{den} y_n)}{\operatorname{den} y_n} \leqq \frac{M}{1} = M < \infty. \tag{4}$$

We infer that

$$\begin{aligned} &\lim_n \operatorname{den} y_n = \infty, \\ &\lim_n f(\operatorname{den} y_n) = 0, \qquad (5) \\ &\lim_n y_n = x. \end{aligned}$$

We let z be such a sequence that, for $n \in \omega$, $z_n = 2 \cdot y_n - x$. Hence for $n \in \omega$, $z_n - x = 2 \cdot (y_n - x)$, $y_n = (z_n + x)/2$, y_n is strictly between z_n and x. Accordingly $\lim_n z_n = x$. Also, for sufficiently large $n \in \omega$, $z_n \in \mathrm{J}$ and

$$\begin{aligned}|g(z_n) - g(x)| &= \left| \sum_{r \in W(z_n)} \frac{f(\operatorname{den} r)}{\operatorname{den} r} - \sum_{r \in W(x)} \frac{f(\operatorname{den} r)}{\operatorname{den} r} \right| \\ &\geqq \frac{f(\operatorname{den} y_n)}{\operatorname{den} y_n} > \lambda \cdot |y_n - x| = \lambda \cdot |z_n - x|/2.\end{aligned}$$

Consequently

$$\overline{\lim_{t \to x}} \frac{g(t) - g(x)}{t - x} \geqq \lambda/2.$$

If x were rational then the multiplication of the first two inequalities in (4) by the positive factor

$$\frac{\operatorname{den} y_n \cdot \operatorname{den} x}{\lambda}$$

and the use of (5) would lead us by well-known reasoning to the absurdity that $1 \leqq 0$.

Application 3.2. Herein we suppose $f(x) = x^{-2}$ whenever $x \in \mathrm{Rp}$. Accordingly $g = w$ and the Liouville number $\sum_{n=0}^{\infty} (1/2)^{n!}$ belongs to A. The invalidity of 1.1 follows from 3.1(1). Moreover, if x is any Liouville number in J, then $x \in B$ and

$$\overline{\lim_{t \to x}} \frac{w(t) - w(x)}{t - x} = \infty.$$

Application 3.3. Turning to Khinchin's Theorem, we let $K = \{x \in \mathrm{R}\colon$ there are infinitely many $r \in \mathrm{F}$ for which $|x - r| < f(\operatorname{den} r)/\operatorname{den} r\}$. Since $\operatorname{den} r = \operatorname{den}(r + n)$ whenever r is rational and n is an integer, we see that K is invariant under integer translation and, from 3.1(2), that the Lebesgue measure of $K \cap \mathrm{J}$ is zero. Consequently, the Lebesgue measure of K is zero and Khinchin's Theorem is at hand.

References

1. R. P. Boas, Jr., A primer of real functions, Carus Mathematical Monographs, Number 13, MAA; Wiley, New York, 1960.

2. M. K. Fort, Jr., A theorem concerning functions discontinuous on a dense set, this MONTHLY, 58 (1951) 408–410.

3. G. A. Heuer, A property of functions discontinuous on a dense set, this MONTHLY, 73 (1966) 378–379.

4. A. Ya. Khinchin, Continued Fractions, 3rd ed., 1961, English Translation, University of Chicago Press, 1964.

5. S. G. Wayment, Sizing up sets and continuity-differentiability relationships, this MONTHLY, 77 (1970) 740–743.

6. David Drasin and Robert Gilmer, Complements and comments, this MONTHLY, 78 (1971) 1104.

PROFESSOR LEO MOSER — REFLECTIONS OF A VISIT

W. E. MIENTKA, University of Nebraska-Lincoln

Professor Leo Moser[1] was known throughout the Mathematical Community as a significant researcher and excellent lecturer.

I first met Leo during the Summer Research Institute in the Theory of Numbers held at the University of Colorado in 1959. After talking with him and hearing his lectures during the Institute, I felt that arrangements would have to be made in the near future for a visit to Nebraska. During the academic year 1962–63 while Professor Moser was on a lecture tour for the MAA, I invited him to present two research lectures to the Nebraska Section on May 3 and 4, 1963. He responded: "Professor D. W. Western of Franklin and Marshall College is my booking agent and I will write him immediately and find out whether it would be possible to clear May 3rd and 4th for me and thus enable me to give the lectures in Nebraska." His generosity was revealed in a subsequent letter in which he asserted: "According to a letter just received from Professor D. W. Western, I am to lecture in Cleveland, Ohio on May 1st and 2nd and in St. Petersburg, Florida on May 6th and 7th. Assuming connections are not too bad I should be able to get to Nebraska in time. If I find that the connections are not easy then I can move the Cleveland date back by one week I imagine. My talks at Nebraska will be on Number Theory and have the general title "Some New Applications of Generating Series."

As usual his lectures were delivered with vigor, humor, and clarity. Following his last lecture I invited him to my office in order to discuss some of his results, and during our conversation the subject of mathematical limericks was mentioned and he asked if I would like to record some of his and other's limericks. (I had previously received his permission to record his lectures.)

The main purpose of this paper is to present a transcription of these limericks and other verse, recorded on May 4, 1963.

Hiawatha Designs an Experiment

Hiawatha, mighty hunter,
He could shoot ten arrows upward,
Shoot them with such strength and swiftness
That the last that left the bull-string
Ere the first to earth descended.

This was commonly regarded
As a feat of skill and cunning.
Several sarcastic spirits
Pointed out to him, however,
That it might be much more useful

[1] Professor Moser died February 9, 1970 at the age of 48. The author wishes to express his appreciation to Mrs. Moser for her permission to publish this paper.

If he sometimes hit the target.
"Why not shoot a little straighter
And employ a smaller sample?"
Hiawatha, who at college
Majored in applied statistics,
Consequently felt entitled
To instruct his fellow man
In any subject whatsoever,
Waxed exceedingly indignant,
Talked about the law of errors,
Talked about truncated normals,
Talked of loss of information,
Talked about his lack of bias,
Pointed out that (in the long run)
Independent observations,
Even though they missed the target,
Had an average point of impact
Very near the spot he aimed at,
With a possible exception
of a set of measure zero.
"This," they said, "was rather doubtful;
Anyway, it didn't matter
What resulted in the long run:
Either he must hit the target
Much more often than at present,
Or himself would have to pay for
All the arrows he had wasted."
Hiawatha, in a temper,
Quoted parts of R. A. Fisher,
Quoted Yates and quoted Finney,
Quoted reams of Oscar Kempthorne,
Quoted Anderson and Bancroft
(practically *in extenso*)
Trying to impress upon them
That what actually mattered
Was to estimate the error.
Several of them admitted:
"Such a thing might have its uses;
Still," they said, "he would do better
If he shot a little straighter."
Hiawatha, to convince them,
Organized a shooting contest.
Laid out in the proper manner
Of designs experimental
Recommended in the textbooks,
Mainly used for tasting tea
(but sometimes used in other cases)
Used factorial arrangements
And the theory of Galois,
Got a nicely balanced layout
And successfully confounded
Second order interactions.
All the other tribal marksmen,
Ignorant benighted creatures
Of experimental setups,
Used their time of preparation
Putting in a lot of practice
Merely shooting at the target.
Thus it happened in the contest
That their scores were most impressive
With but one solitary exception.
This, I hate to have to say it,
Was the score of Hiawatha,
Who as usual shot his arrows,
Shot them with great strength and swiftness,
Managing to be unbiased,
Not however with a salvo
Managing to hit the target.
"There!" they said to Hiawatha,
"That is what we all expected."
Hiawatha, nothing daunted,
Called for pen and called for paper.
But analysis of variance
Finally produced the figures
Showing beyond all peradventure,
Everybody else was biased.
And the variance components
Did not differ from each other's,
Or from Hiawatha's.
(This last point it might be mentioned,
Would have been much more convincing
If he hadn't been compelled to

Estimate his own components
From experimental plots on
Which the values all were missing.)
Still they couldn't understand it,
So they couldn't raise objections.
(Which is what so often happens
with analysis of variance.)
All the same his fellow tribesmen,
Ignorant benighted heathens,
Took away his bow and arrows,
Said that though my Hiawatha
Was a brilliant statistician,
He was useless as a bowman.
As for variance components
Several of the more outspoken
Made primeval observations
Hurtful of the finer feelings
Even of the statistician.
In a corner of the forest
Sits alone my Hiawatha
Permanently cogitating
On the normal law of errors.
Wondering in idle moments
If perhaps increased precision
Might perhaps be sometimes better
Even at the cost of bias,
If one could thereby now and then
Register upon a target.

* * *

Chicago's mathematical forces
Despite their numerous resources
Always adorn
With the Lemma of Zorn
At least ninety percent of their courses.

* * *

Professor Adrian Albert said who
Can tell me a theorem that's true
The ones that I know
Are simply not so
When the characteristic is two.

* * *

Eduard Čech by God's grace
Was the first man on Earth to trace
That sordid and dreary
Cohomology theory
Of a subnormal bicompact space.

* * *

A mathematician confided
That a Möbius strip is one sided
And you get quite a laugh
When you cut it in half
Because it stays in one piece when divided.

* * *

Mathematicians try hard to floor us
With a non-orientable torus
The bottle of Klein
They say is divine
But it is so exceedingly porous.

* * *

Once a man whose name wouldn't rhyme
Found an unbelievably large prime
But with no place to store it
He had no use for it
So Dick Lehmer got it for a dime.

* * *

A mathematician named Moser
Well-known as a problem proposer
Sent some that were silly
To his brother named Willy
Could he stump him? The answer is no, sir.

* * *

There was a young man from Racine
Who invented a brain-like machine
It knew digits in π
And found cube roots of i
And sang a few hymns in between.

* * *

Where are the zeroes of zeta of s?
Bernhard Riemann made a pretty good guess:
"They're all on the critical line," said he
"And their density is t over $2\pi \log t$."

Now the statement of Riemann has set off a trigger,
And many a good man with vim and with vigor
Tried to prove with mathematical rigor
What happens to zeta as mod t gets bigger.

The names of Hardy, Landau, and Cramér
And Littlewood and Titchmarsh are there.
But in spite of their skill and in spite of finesse
In locating the zeros, no-one's had success.

In 1914, G. H. Hardy did find
An infinite number that lay on the line.
But unfortunately his theorem won't rule out the case
That there may be some zeros in some other place.

Oh where are the zeros of zeta of s?
We must know exactly, we cannot just guess.
For in order to refine the prime number theorem,
The path of integration must not get too near 'em.

(by Tom Apostol*)

* * *

There was a young fellow named Ben
Who could only count modulo ten
He said when I go
Past my last little toe
I shall have to start over again.

* * *

The binary system is fun
For with it strange things can be done
And two as you know
Is a one and an oh
And five is one hundred and one.

* * *

The marvelous things a computer can do
Makes an idiot out of the highest IQ
But there's one consolation
In this observation
It can't even add up to two.

* * *

Here's to uncle Albert E.
Pundit of relativity
You'll know him by his fiddler's locks
and by his utter lack of socks.

Here's to uncle Oswald V.
Lover of England and her tea
He is that mathematician of note
Who needs four buttons to button his coat.

* * *

Condemned for defending the masses
Scourged for defaming the lasses
Not moved by disgrace
He has come to this place
To teach the class of all classes.
(Student – University of Minnesota,
written on the occasion of B. Russell's
visit in 1942–1943)

* * *

A function from feeling inferior
Felt life monotonically drearier
With a hell of a yell
That jumped into L
And converged to the limit superior.

There once was a hairy baboon
Who always breathed down a bassoon
For he said it appears
That in millions of years
I will certainly hit on a tune.

* * *

Let x stand for beauty, y manners well-bred,
Zed fortune (this last is essential).
Let L stand for love, our philosophers said,
Then L is a function of x, y, and zed
Of the kind that is known as potential.
Now integrate L with respect to dt
(t standing for time and persuasion).
Then between proper limits it's easy to see
The definite integral marriage must be.
(A. S. Eddington — A very concise demonstration. L.M.)

* * *

Said a monkey as he swung by his tail
To his children both female and male
From your offsprings my dears
In some millions of years
May emerge a professor at Yale.

But who could dream in those
times immemorial
That from those creatures arboreal
Professor Uhler would evolve
Who had the courage and resolve
To calculate one thousand factorial.

* * *

Nature and nature's law lay hid by night
Then God said "Let Newton be,"
and all was light,
This could not last, the Devil shouting "Ho
Let Einstein be," restored the status quo.

* * *

A mathematician named Moser
Was able to regain his composure
When a pair of young men
Claimed their distance was ten
But were unable to prove this disclosure.

* * *

A mathematician O'Flaherty
Invented a new singularity
Where the Z plane corrodes
And the function explodes
Well you'll have to admit it's a rarity.

* * *

The subject of today's instruction
Is to perform mathematical induction.
The steps are easy as one two three
If you want to get clued just listen to me.

You want to prove a theorem then
To prove it for every integer n
You prove it first for n equal one
And then the induction is begun.

The proof goes on in the following way
You assume it next for n equal k.
If you can show it then for $k + 1$
Then the induction is truly done.
(To be sung to the tune of
"Three Little Fishes")

* * *

* Prof. Apostol points out that the oral tradition has produced some changes in his verses. He offers the original, guaranteed correct, version of what turns out to be a *song*, sung to the tune of "Sweet Betsy from Pike". Our efforts to locate the melody have failed. *Editor.*

Where are the zeros of zeta of s?

Where are the zeros of zeta of s?
G. F. B. Riemann has made a good guess,
They're all on the critical line, said he,
And their density's one over $2\pi \log t$.

This statement of Riemann's has been like a trigger,
And many good men, with vim and with vigor,
Have attempted to find, with mathematical rigor,
What happens to zeta as mod t gets bigger.

The names of Landau and Bohr and Cramér,
And Hardy and Littlewood and Titchmarsh are there,
In spite of their efforts and skill and finesse,
In locating the zeros no one's had success.

In 1914 G. H. Hardy did find,
An infinite number that lay on the line,
His theorem, however, won't rule out the case,
That there might be a zero at some other place.

Let P be the function π minus li,
The order of P is not known for x high,
If square root of x times log x we could show,
Then Riemann's conjecture would surely be so.

Related to this is another enigma,
Concerning the Lindelöf function $\mu(\sigma)$
Which measures the growth in the critical strip,
And on the number of zeros it gives us a grip.

But nobody knows how this function behaves,
Convexity tells us it can have no waves,
Lindelöf said that the shape of its graph,
Is constant when sigma is more than one-half.

Oh, where are the zeros of zeta of s?
We must know exactly,we cannot just guess,
In order to strengthen the prime-number theorem,
The path of integration must not get too near 'em.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

THE LOGARITHMIC MEAN

B. C. CARLSON, Iowa State University

Let the *logarithmic mean* of the positive numbers x and y be defined by

$$(1)\qquad \begin{aligned} L(x,y) &= \frac{x-y}{\log x - \log y}, \quad x \neq y, \\ L(x,x) &= x. \end{aligned}$$

Note that L is symmetric and homogeneous in x and y and continuous at $x = y$. It is not widely known that L separates the arithmetic and geometric means:

$$(2)\qquad (xy)^{\frac{1}{2}} \leqq L(x,y) \leqq \frac{x+y}{2},$$

with strict inequalities if $x \neq y$. Division by y shows that (2) is equivalent to well-known inequalities in the single variable $w = x/y$, but the beauty of (2) comes from its symmetry in two variables. The right-hand inequality is due to Ostle and Terwilliger [**1**], and several proofs are cited by Mitrinović [**2**]. In both sources the symmetry is somewhat slighted by retaining the unnecessary condition $x \geqq y$. The left-hand inequality was stated by Carlson [**3**, Eq. (3.1)], who obtained (2) by specializing some rather general integral inequalities to the case of the representation

$$(3)\qquad \frac{1}{L(x,y)} = \int_0^1 \frac{du}{ux + (1-u)y}.$$

In the present note we first prove and sharpen (2) by an elementary method which treats x and y symmetrically.

THEOREM 1. *If the positive numbers x and y are unequal, then*

$$(4)\qquad (xy)^{\frac{1}{2}} < (xy)^{\frac{1}{4}}\frac{\sqrt{x}+\sqrt{y}}{2} < L(x,y) < \left(\frac{\sqrt{x}+\sqrt{y}}{2}\right)^2 < \frac{x+y}{2}.$$

Proof. If $t > 0$ the inequality of the arithmetic and geometric means implies that

$$t^2 + t(x+y) + \left(\frac{x+y}{2}\right)^2 > t^2 + t(x+y) + xy > t^2 + 2t(xy)^{\frac{1}{2}} + xy.$$

Thus

$$\int_0^\infty \frac{dt}{\left(t+\frac{x+y}{2}\right)^2} < \int_0^\infty \frac{dt}{(t+x)(t+y)} < \int_0^\infty \frac{dt}{(t+\sqrt{xy})^2}.$$

Evaluating the middle integral by the method of partial fractions, we find

$$\frac{2}{x+y} < \frac{1}{x-y} \lim_{R\to\infty} [\log(t+y) - \log(t+x)]_0^R < \frac{1}{\sqrt{xy}},$$

which implies (2). We now sharpen (2) by replacing x by $\sqrt{x}$ and y by $\sqrt{y}$:

$$(xy)^{\frac{1}{4}} < \frac{2(\sqrt{x}-\sqrt{y})}{\log x - \log y} < \frac{\sqrt{x}+\sqrt{y}}{2}.$$

Multiplication by $(\sqrt{x}+\sqrt{y})/2$ proves the two inner inequalities in (4). The two outer ones follow from the inequality of the arithmetic and geometric means.

The process by which (2) was sharpened can be repeated to obtain (8). Instead of taking this route we prove a more general inequality first. For any real $t \neq 0$ and any positive x and y, we define

$$G_t(x,y) = t(xy)^{t/2}\frac{x-y}{x^t-y^t}, \quad A_t(x,y) = t\frac{x^t+y^t}{2}\frac{x-y}{x^t-y^t}, \quad x \neq y, \tag{5}$$
$$G_t(x,x) = A_t(x,x) = x.$$

If we further define $G_0(x,y) = A_0(x,y) = L(x,y)$, it is easy to verify that G_t and A_t are continuous in t. They are also positive and even in t.

THEOREM 2. *If x and y are positive and t is real, then*

$$G_t(x,y) < L(x,y) < A_t(x,y), \qquad t(x-y) \neq 0. \tag{6}$$

The first and third members are respectively decreasing and increasing functions of $|t|$, and the sharpness of the inequalities is measured by

$$A_t^2(x,y) - G_t^2(x,y) = \tfrac{1}{4}t^2(x-y)^2. \tag{7}$$

Proof. In (2) replace x by x^t and y by y^t and multiply by the positive quantity $t(x-y)/(x^t-y^t)$ to get (6). By straightforward calculation,

$$t\frac{dG_t}{dt} = G_t\left(1-\frac{A_t}{L}\right), \quad t\frac{dA_t}{dt} = A_t - \frac{G_t^2}{L},$$

from which it follows by (6) that A_t increases with $|t|$ while G_t decreases. Incidentally, a second differentiation shows that A_t is convex and $1/G_t$ is log convex in t.

COROLLARY 1. *If x and y are positive and unequal and n is a nonnegative*

integer, then

$$(xy)^{2^{-n-1}} \prod_{m=1}^{n} \alpha_m(x,y) < L(x,y) < \alpha_n(x,y) \prod_{m=1}^{n} \alpha_m(x,y), \tag{8}$$

where

$$\alpha_m(x,y) = \frac{x^{2^{-m}} + y^{2^{-m}}}{2}.$$

The products are taken to be unity if $n = 0$. The first and third members of (8) *are respectively increasing and decreasing functions of n, and the difference of their squares is $2^{-2n-2}(x-y)^2$.*

Proof. Choose $t = 2^{-n}$ in Theorem 2 and note that

$$x - y = (x^{2^{-n}} - y^{2^{-n}}) \prod_{m=1}^{n} (x^{2^{-m}} + y^{2^{-m}}).$$

The inequalities (8) reduce to (2) if $n = 0$ and to the inner inequalities of (4) if $n = 1$. As $n \to \infty$ we obtain the following infinite product.

COROLLARY 2. *If x and y are positive numbers, then*

$$L(x,y) = \prod_{m=1}^{\infty} \alpha_m(x,y). \tag{9}$$

An equality or inequality for $L(x,y)$ of course implies a corresponding result for $\log x$ obtained by putting $y = 1$. For example, (6) gives

$$\frac{2}{t}\,\frac{x^t - 1}{x^t + 1} < \log x < \frac{x^t - 1}{t x^{t/2}}, \quad t \neq 0,\ x > 1, \tag{10}$$

with reversed inequalities if $0 < x < 1$. The inequalities become sharper as $|t|$ decreases and as $|x - 1|$ decreases. Likewise (9) implies, for $x > 0$,

$$\log x = (x - 1) \prod_{m=1}^{\infty} \frac{2}{1 + x^{2^{-m}}}. \tag{11}$$

Finally we give an algorithm for computing $L(x,y)$ or $\log x$ by recurrence relations. As $t \to 0$ we find by developing (5) in powers of t that A_t and G_t differ from L by terms of order t^2, but

$$L(x,y) = \tfrac{1}{3}\{A_t(x,y) + 2G_t(x,y)\}\,\{1 + \delta_t(x,y)\}, \tag{12}$$

where δ_t is of order t^4. Since

$$A_{t/2} = \tfrac{1}{2}(A_t + G_t), \quad G_{t/2} = (A_{t/2} G_t)^{\frac{1}{2}}, \tag{13}$$

extraction of one square root cuts t in half and ultimately reduces the fractional

error δ_t by a factor of 16. For small t it is difficult to calculate A_t and G_t directly from (5) owing to cancellation in $x^t - y^t$, but use of (13) avoids this problem. We define $a_n = A_t$ and $g_n = G_t$, where $t = 2^{1-n}$, and proceed as follows.

ALGORITHM. *If x and y are positive numbers, let*

$$
\begin{aligned}
a_1 &= \tfrac{1}{2}(x+y), & g_1 &= (xy)^{\frac{1}{2}}, \\
a_{n+1} &= \tfrac{1}{2}(a_n + g_n), & g_{n+1} &= (a_{n+1}g_n)^{\frac{1}{2}}, \quad n = 1,2,3,\cdots.
\end{aligned} \tag{14}
$$

Then the common limit of a_n and g_n as $n \to \infty$ is the logarithmic mean $L = L(x,y)$ defined by (1). *Moreover,*

$$L = \tfrac{1}{3}(a_n + 2g_n)(1+\varepsilon_n)^{-1}, \tag{15}$$

where

$$0 \leqq \varepsilon_n \leqq \frac{2^{-4n}}{180}\left(\frac{x-y}{g_n}\right)^4 \leqq \frac{2^{-4n-2}(x-y)^4}{45x^2y^2}. \tag{16}$$

The recurrence relations (14) are those of Borchardt's algorithm [4]. We omit the proof of the error bounds (16) by expansion in power series, because a method of further speeding the convergence will be discussed elsewhere [5]. An algorithm with slower convergence is given in [4, Eq. (2.4)].

Note added in proof. Corollary 1 provides a solution of the second of two problems proposed by D. S. Mitrinović, Problem 5626, this MONTHLY, 75 (1968) 911–912. See also [2, pp. 383–384].

This work was performed in the Ames Laboratory of the U. S. Atomic Energy Commission,

References

1. B. Ostle and H. L. Terwilliger, A comparison of two means, Proc. Montana Acad. Sci., 17 (1957) 69–70.

2. D. S. Mitrinović, Analytic Inequalities, Springer-Verlag, Berlin, 1970, p. 273.

3. B. C. Carlson, Some inequalities for hypergeometric functions, Proc. Amer. Math. Soc., 17 (1966) 32–39.

4. ———, Algorithms involving arithmetic and geometric means, this MONTHLY, 78 (1971) 496–505.

5. ———, An algorithm for computing logarithms and arctangents, Math. Comp., April 1972.

ON THE CONVERGENCE OF THE L^p NORM TO THE L^∞ NORM

R. A. HANDELSMAN, University of Illinois and
J. S. LEW, IBM Thomas J. Watson Research Center

If $f(x)$ is a real- or complex-valued Lebesgue-measurable function on the finite or infinite interval $[a,b]$, then its L^p norm is defined by

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx\right)^{1/p}$$

for $1 \leqq p < \infty$, and its L^∞ norm is defined by

$$\|f\|_\infty = \text{ess. sup. } \{|f(x)| : a \leqq x \leqq b\}$$

with respect to Lebesgue measure. We admit ∞ as a possible value for these norms, so that $\|f\|_p$ is well-defined for $1 \leqq p \leqq \infty$, and we recall the following well-known result of integration theory.

THEOREM 1. *If $f(x)$ is Lebesgue-measurable on $[a, b]$ and $\|f\|_q$ is finite for some $q < \infty$, then*

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty \tag{1}$$

Proof: [**4**, p. 39].

In this note we shall amplify the result just stated by using a basic tool of asymptotic analysis, a standard Abelian theorem for Laplace transforms which is usually called Watson's Lemma. To do this we must impose stronger assumptions on $f(x)$ (which are justified in many practical situations), but in return we can exhibit not merely an alternative proof of (1) but additionally an estimate for the resulting order of convergence. Moreover, we obtain as a by-product the Stirling approximation for p! We recall first the statement of this lemma.

WATSON'S LEMMA. *Let $g(t)$ be a locally integrable function on $[0, \infty)$ which is $O[\exp(rt)]$ as $t \to \infty$ for some finite constant r, and which has a finite asymptotic expansion*

$$g(t) = \sum_{j=0}^{n} a_j t^{c_j - 1} + o(t^{c_n - 1}) \quad \text{as } t \to 0+ \tag{2}$$

with $0 < \operatorname{Re}(c_0) < \cdots < \operatorname{Re}(c_n)$. Then the integral

$$I(s) = \int_0^\infty \exp(-st) g(t)\, dt \tag{3}$$

has a finite asymptotic expansion

$$I(s) = \sum_{j=0}^{n} a_j \Gamma(c_j) s^{-c_j} + o(s^{-c_n}) \quad \text{as } s \to +\infty, \tag{4}$$

where $\Gamma(c)$ denotes the usual gamma function.

Proof: [**1**, pp. 49–50] or [**2**, pp. 31–34]; for a generalization see [**3**].

Equation (4) can be obtained formally through substituting (2) into (3) and integrating term by term, so that Watson's Lemma simply validates this formal computation. Using this lemma, we now prove a result which complements Theorem 1 when $\|f\|_\infty$ is finite. In our treatment we may assume $f(x) \geqq 0$, since $f(x)$ and $|f(x)|$ have the same L^p norms.

THEOREM 2. *Let $f(x)$ be a nonnegative Lebesgue-measurable function on the finite or infinite interval $[a, b]$, and let $\|f\|_q$ be finite for some $q < \infty$. Suppose also for some $r > a$ that $f(x)$ is continuous and positive on $[a, r]$ while $f'(x)$ is defined, continuous, and negative on (a, r); and that*

$$f(a) > M = \text{ess. sup}\{f(x): r \leqq x \leqq b\}, \tag{5}$$

whence $f(a)$ is the unique maximum of $f(x)$ on $[a, b]$. Suppose finally for some c, $k > 0$ that

$$f(a) - f(x) = k(x-a)^{1/c} + o[(x-a)^{1/c}] \text{ as } x \to a+ \tag{6}$$

and that (6) *may be differentiated to yield*

$$f'(x) = -kc^{-1}(x-a)^{-1+1/c} + o[(x-a)^{-1+1/c}] \text{ as } x \to a+. \tag{7}$$

Then as $p \to \infty$,

$$\|f\|_p = f(a)[1 - cp^{-1}\log p + p^{-1}\log C + o(p^{-1})], \tag{8}$$

where C is given by

$$C = c\Gamma(c)[f(a)/k]^c. \tag{9}$$

Proof. To estimate $\|f\|_p$ we write

$$(\|f\|_p)^p = I_1(p) + I_2(p) = \int_a^r f(x)^p dx + \int_r^b f(x)^p dx$$

and consider $I_1(p)$ as $p \to \infty$. If we introduce a new variable t through $f(x) = f(a)\exp(-t)$ and a new function $g(t)$ through

$$g(t) = \begin{cases} dx/dt \text{ on } [0, \log(f(a)/f(r))], \\ 0 \text{ elsewhere on } [0, \infty), \end{cases}$$

then $g(t)$ is well-defined and locally integrable on $[0, \infty)$, so that $I_1(p)$ takes the form

$$I_1(p) = f(a)^p \int_0^\infty \exp(-pt)g(t)\,dt.$$

Moreover, by computation with (6) and (7) we find

$$g(t) = c[f(a)/k]^c t^{c-1} + o(t^{c-1}) \quad \text{as } t \to 0+$$

so that, by Watson's Lemma and definition (9), we obtain

$$I_1(p) = f(a)^p[Cp^{-c} + o(p^{-c})] \quad \text{as } p \to \infty. \tag{10}$$

However, in $[r, b]$ clearly $f(x)/M \leqq 1$ by assumption (5), so that for all $p \geqq q$

$$I_2(p) = M^p \int_r^b [f(x)/M]^p dx \leqq M^p \int_r^b [f(x)/M]^q dx = KM^p \tag{11}$$

with K independent of p, and thus

$$\|f\|_p = f(a)\exp[p^{-1}(\log C - c\log p)] \cdot [1 + o(1) + O(M/f(a))^p]^{1/p} \tag{12}$$

as $p \to \infty$. By (5) we observe finally that the O-term in (12) is exponentially small as $p \to \infty$, so that we can expand (12) to first order and recover (8).

The first nontrivial term in our assumed form (6) for $f(x)$ has order $1/c$, but nevertheless the first three terms in the derived series (8) for $\|f\|_p$ have orders independent of c; since these terms arise through expanding the exponential factor in (12), which comes from nothing more than the first term in (10). However if we explicitly assume a more elaborate expansion for $f(x)$, then we can explicitly calculate some higher terms in the series for $\|f\|_p$; and indeed the resulting orders of such terms will reflect the spacing of exponents in our assumed series for $f(x)$.

We can easily extend Theorem 2 and allow $f(x)$ to achieve its maximum at any finite number of points in $[a, b]$; since then we need only decompose $[a, b]$ into subintervals each of which has exactly one of these maxima at either its left or right endpoint. In this generalization the smoothness conditions assumed in Theorem 2 near the point a must clearly be imposed in a small half-neighborhood on either side of each maximum. Indeed, to illustrate this remark simply, let us consider a nonnegative continuous $f(x)$ on $[a, b]$ which has a unique maximum at v in (a, b) and is C^2 in some neighborhood of this v. Then by Taylor's theorem

$$f(x) = f(v) + \tfrac{1}{2}f''(v)(x - v)^2 + o[(x - v)^2] \quad \text{as } x \to v+$$

so that we may write

$$(\|f\|_p)^p = \sum_{j=1}^{4} I_j(p) = \left[\int_a^u + \int_u^v + \int_v^w + \int_w^b\right] f(x)^p dx \tag{13}$$

for some u and w sufficiently near v, and we may proceed as in Theorem 2.

As in (11) and (12) we may argue that $I_1(p)$ and $I_4(p)$ are exponentially small, so that we need only approximate $I_2(p)$ and $I_3(p)$. However by (10) in Theorem 2 we already have

$$I_3(p) = f(v)^p[Cp^{-\frac{1}{2}} + o(p^{-\frac{1}{2}})] \quad \text{as } p \to \infty, \tag{14}$$

where, by definition (9), we now have

$$C = \tfrac{1}{2}\left|2\pi f(v)/f''(v)\right|^{\frac{1}{2}}. \tag{15}$$

Moreover, to find $I_2(p)$ we may either repeat the argument of Theorem 2 or simply change x into $-x$; and by either method we obtain

$$I_2(P) = f(v)^p[Cp^{-\frac{1}{2}} + o(p^{-\frac{1}{2}})] \quad \text{as } p \to \infty \tag{16}$$

so that $I_2(p)$ and $I_3(p)$ coincide to lowest order. If we substitute (14) and (16) into (13), and take the p'-th root of (13) as before, then finally we obtain

$$\|f\|_p = f(v)[1 - \tfrac{1}{2}p^{-1}\log p + \tfrac{1}{2}p^{-1}\log|\pi f(v)/f''(v)| + o(p^{-1})]$$

as $p \to \infty$. In this not unrepresentative case, the convergence to $\|f\|_\infty$ is obviously slow, so that $\|f\|_\infty$ is often a poor estimate for $\|f\|_p$.

As a further specialization let $f(x) = x\exp(-x)$ on $[0, +\infty)$, so that $f(x)$ has a unique maximum at $x = 1$ and a convergent expansion

$$f(x) = e^{-1}[1 - \tfrac{1}{2}(x-1)^2 + \tfrac{1}{3}(x-1)^3 + \cdots].$$

Then $(\|f\|_p)^p$ can be calculated exactly in terms of the gamma function, and approximately by means of our results; indeed by (13)–(16)

$$\Gamma(1+p) = (\|f\|_p)^p p^{1+p} \sim 2Cp^{-\frac{1}{2}}f(1)^p p^{1+p} = (2\pi p)^{\frac{1}{2}}p^p e^{-p}.$$

This is the Stirling approximation for $p!$, which may, of course, be obtained directly through Watson's lemma.

The initial work leading to this paper took place while both authors were staff members of the Division of Applied Mathematics at Brown University, at which time this work was supported by the Advanced Research Projects Agency, Department of Defense, under ARPA Contract SD–86 with Brown University.

References

1. E. T. Copson, Asymptotic Expansions, Cambridge Univ. Press, Cambridge, 1965.
2. A. Erdelyi, Asymptotic Expansions, Dover, New York, 1956.
3. R. A. Handelsman and J. S. Lew, Asymptotic expansion of a class of integral transforms via Mellin transforms, Arch. Rational Mech. Anal, 35 (1969) 382–396.
4. L. H. Loomis, An Introduction to Abstract Harmonic Analysis, Van Nostrand, Princeton, N. J., 1951.

EXTENSION OF MAPPINGS IN FINITE ABELIAN GROUPS

Kyle D. Wallace, Western Kentucky University

The purpose of this note is to provide a simple example to resolve a question that may arise in an introductory course in abstract algebra, and further to provide a means for injecting more elementary structure theory in such a course. Consideration shall thus be restricted to finite abelian groups (written additively). If A is a subgroup of the abelian group G and ϕ is an automorphism of G, then ϕ induces isomorphisms which yield $A \cong \phi(A)$ and $G/A \cong G/\phi(A)$. In fact, the diagram

$$\begin{tikzcd} A \arrow[r, tail] \arrow[d, tail, two heads] & G \arrow[r, two heads] \arrow[d, tail, two heads] & G/A \arrow[d, tail, two heads] \\ \phi(A) \arrow[r, tail] & G \arrow[r, two heads] & G/\phi(A) \end{tikzcd}$$

is commutative, where the vertical maps are induced by ϕ and the horizontal maps are canonical. The problem to be considered deals with the converse of this result.

QUESTION. *Suppose A and B are subgroups of the finite abelian group G such that $A \cong B$ and $G/A \cong G/B$. Does there exist an automorphism α of G such that $\alpha(A) = B$?*

We note that it suffices to consider only finite abelian *p-groups* for p a prime. Before giving an example to illustrate that such an automorphism need not exist, we shall require the concept of the height of an element in a p-group and a few elementary properties.

Let G be an abelian p-group. We define inductively: $p^0G=G$, $pG=\{px \mid x \in G\}$, $\cdots, p^{n+1}G = p(p^nG)$. For $x \in G$, the height of x in G, $h_G(x)$, is the non-negative integer n if $x \in p^nG$ but $x \notin p^{n+1}G$ (that is, if x is divisible by p^n but not by p^{n+1}). We say that x has infinite height in G if $x \in p^nG$ for each positive integer n. Note that for a finite p-group G, 0 is the only element of infinite height. We write $h_G(0) = \infty$ and $\infty > n$ for each positive integer n. If x lies in a subgroup A of G, we may define two heights for x; namely $h_A(x)$ and $h_G(x)$ the height of x in A and G, respectively. The following properties are easily established.

P1: (a) If $h_G(x) \neq h_G(y)$ then $h_G(x + y) = \min\{h_G(x), h_G(y)\}$.
(b) If $h_G(x) = h_G(y)$ then $h_G(x + y) \geqq h_G(x)$, and may be strictly larger.

P2: If G is a direct sum of subgroups, then height is computed componentwise. (That is, if $G = \oplus_{i\in I} A_i$, $g = \sum_{i \in I} a_i$ then $h_G(g) = \min\{h_{A_i}(a_i)\}$).

P3: (a) If f is a homomorphism of G into $\bar{G}$ then $h_G(x) \leqq h_{\bar{G}}(f(x))$ for $x \in G$.
(b) An automorphism preserves height.

Example. Let $G = \langle a\rangle \oplus \langle b\rangle \oplus \langle c\rangle$ be a direct sum of cyclic groups with

$$\langle a\rangle \cong Z_p,\ \langle b\rangle \cong Z_{p^2},\ \text{and}\ \langle c\rangle \cong Z_{p^3}.$$

Let $A = \langle a\rangle \oplus \langle pc\rangle$ and $B = \langle b\rangle \oplus \langle p^2c\rangle$. Then $A \cong B \cong G/A \cong G/B \cong Z_p \oplus Z_{p^2}$, and the conditions of the question are satisfied. Note that the element a has height 0 and order p in G. Any element x of order p in B may be represented in the form $x = mpb + np^2c$ and hence $h_G(x) \geqq 1$. Since an automorphism of G must preserve both height and order, we have $\alpha(a) \notin B$ for any automorphism α of G.

The desired automorphism could not exist in the above example due to the fact that no isomorphism from the subgroup A onto the subgroup B could preserve height as computed in G. As a corollary to one of the nicest structure theorems in mathematics, (see Paul Hill's generalization of Ulm's theorem [2], or [3]), it follows that for finite abelian groups this condition of preservation of height is sufficient.

THEOREM. *Let A, B be subgroups of the finite abelian p-group G such that $G/A \cong G/B$. If ϕ is an isomorphism of A onto B such that, for each $a \in A$, $h_G(a) = h_G(\phi(a))$ then ϕ can be extended to an automorphism of G.*

For finite abelian p-groups, it seems feasible to ask whether a sufficiency condition can be obtained by replacing the condition of preservation of height by a

condition on the subgroup A. Such a condition follows from the following proposition.

PROPOSITION. *Let A be a direct summand of the finite p-group G. If $A \cong B$ and $G/A \cong G/B$ then B is a direct summand of G.*

To prove this proposition, we need one further preliminary result. The subgroup A is a *pure* subgroup of the p-group G if $p^nA = A \cap p^nG$ for each positive integer n. Note $p^nA \subseteq A \cap p^nG$. Clearly if A is a direct summand of G then A is pure in G. A pure subgroup need not be a direct summand. However, in certain special cases pure subgroups are direct summands (see [1] and Theorems 5 and 7 in [4]). In particular, suppose that A is a pure subgroup of the finite abelian p-group G. Then G/A is finite and hence a direct sum of cyclic groups, say

$$G/A = \bigoplus_{i=1}^{n} \langle x_i + A \rangle .$$

Since A is pure in G, there exists for each $i = 1, \cdots, n$, an element $y_i \in G$ such that $y_i + A = x_i + A$ and $\mathcal{O}_G(y_i) = \mathcal{O}_{G/A}(x_i + A)$. Let B be the subgroup of G generated by the elements y_i, $i = 1, \cdots, n$. Then it follows that $G = A \oplus B$. Hence, if A is a pure subgroup of the finite abelian p-group G, then A is a direct summand of G.

Proof of Proposition. By the above discussion, it suffices to show that B is a pure subgroup of G. Since A is a direct summand of G, $G \cong A \oplus G/A$. Thus for each positive integer n,

$$\begin{aligned} p^nG \cong p^n(A \oplus G/A) &= p^nA \oplus p^n(G/A) \cong p^nB \oplus p^n(G/B) \\ &= p^nB + (p^nG + B/B) \cong p^nB + (p^nG/B \cap p^nG). \end{aligned}$$

Consequently, it follows that the order of p^nB is equal to the order of $B \cap p^nG$. Since $p^nB \subseteq B \cap p^nG$, $p^nB = B \cap p^nG$ for each positive integer n and B is pure in G.

COROLLARY. *Let A and B be finite abelian groups and ϕ a monomorphism of A into $A \oplus B$. If*

$$A \overset{\phi}{\rightarrowtail} A \oplus B \twoheadrightarrow B$$

is an exact sequence then it is split exact.

References

1. L. Fuchs, A. Kertesz, and T. Szele, Abelian groups in which every serving subgroup is a direct summand, Publications Math. Debrecen, 3 (1953) 95–105.

2. P. Griffith, Infinite Abelian Group Theory, The University of Chicago Press, Chicago, 1970.

3. P. Hill, On the classification of abelian groups, (Preprint).

4. I. Kaplansky, Infinite Abelian Groups, University of Michigan Press, Ann Arbor, 1954.

A PROOF OF GANDHI'S FORMULA FOR THE nth PRIME

CHARLES VANDEN EYNDEN, Illinois State University

Let Q denote the product of the primes less than the odd prime p, and let μ be the Möbius function. The inequalities

$$1 < 2^p\left(-\frac{1}{2}+\sum_{d\mid Q}\frac{\mu(d)}{2^d-1}\right) < 2,$$

which were announced by J. M. Gandhi at the August 1966, International Mathematics Conference in Moscow, allow p to be calculated from the primes preceding it, since for any real number α the inequalities $1 < 2^k\alpha < 2$ hold for at most one integer k. Gandhi's proof involved generating functions; in this note I present a more elementary argument.

Denote the summation by σ. Then

$$(2^Q-1)\sigma = \sum_{d\mid Q}\mu(d)\frac{2^Q-1}{2^d-1} = \sum_{d\mid Q}\mu(d)(1+2^d+2^{2d}+\cdots+2^{Q-d}).$$

Since for $0 \leqq t < Q$ a term $\mu(d)2^t$ occurs each time d is a common divisor of Q and t, the coefficient of 2^t in the last sum is

$$\sum_{d\mid(t,Q)}\mu(d).$$

But it is well known that this is 1 when $(t,Q) = 1$ and 0 otherwise [1]. Thus

$$\sigma = \frac{1}{2^Q-1}\sum_{0<t<Q}^{*} 2^t,$$

where the star indicates that t is restricted to integers prime to Q. Note that $Q-1$ is the largest such t. Then

$$-\tfrac{1}{2}+\sigma = \frac{-(2^Q-1)+\sum^{*}_{0<t<Q} 2^{t+1}}{2(2^Q-1)} = \frac{1+\sum^{*}_{0<t<Q-1} 2^{t+1}}{2(2^Q-1)}.$$

For $2 \leqq j < p$ some prime smaller than p divides $Q-j$; so the largest t occurring in the last summation is $Q-p$. From this it is easy to show that

$$\frac{2^{Q-p+1}}{2\cdot 2^Q} < -\frac{1}{2}+\sigma < \frac{2^{Q-p+2}}{2\cdot 2^Q}.$$

The result follows from multiplication by 2^p.

Reference

1. Ivan Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, 2nd ed., Wiley, New York, 1966, p. 96.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

THE HADAMARD MAXIMUM DETERMINANT PROBLEM

JOEL BRENNER, University of Arizona and LARRY CUMMINGS, University of Waterloo

In 1893 Jacques Hadamard published his classic proof [**12**] that any complex $n \times n$ matrix A with entries in the unit disc satisfies

$$|\det A| \leqq n^{n/2}. \tag{1}$$

Equality is always attained by the Vandermonde of the nth roots of unity [**8**, p. 331]. If the entries of A are restricted to be real, Hadamard remarked that a necessary condition for equality in (1) is $n = 1,2$ or $n \equiv 0 \pmod 4$. Real matrices with determinant $n^{n/2}$ are now appropriately called Hadamard matrices and Hadamard conjectured the existence of a Hadamard matrix for every positive multiple of 4. This still unresolved question has attracted a great deal of effort. Less attention has been devoted to the equally challenging problem of determining the maximum value $|\det A|$ can attain when n is not a multiple of 4. (Since the case $n = 4k$ is surveyed for real entries in [**14**], this note pays detailed attention to the other cases.) A related problem would restrict the entries A to a sector $|\theta| < \theta_0$ of $r = 1$.

The maximum determinant problem is of interest in several diverse areas of mathematics. In statistics it arises in the theory of weighing designs [**18, 19, 20**]. It appears in the study of the rate of convergence of Fredholm expansions for certain types of kernels [**10**]. Combinatorial applications include (v, k, λ) configurations [**21**].

If the entries of A are bounded in modulus by an arbitrary real number M then (1) becomes

$$|\det A| \leqq M^n n^{n/2} \tag{2}$$

[**8**; problem 522] and in case the entries are real

$$|\det A| \leqq M^n 2^{-n}(n+1)^{(n+1)/2}$$

holds [**8**; problem 523].

Consider any real $n \times n$ matrix $A = (a_{ij})$ with $|a_{ij}| \leqq 1$. Expanding the determinant of any such A by minors along successive rows it is apparent that det A is dominated by the determinant of a $(-1, 1)$ matrix; i.e., a matrix all of whose entries

are either -1 or 1. Since there are finitely many such matrices the maximum determinant problem has a solution for each n. There are two questions here: the computation of the maximum value α_n of the determinant for each n and the determination of those classes of $(-1,1)$ matrices whose determinants attain the maximum value.

For n odd G. Barba [1] in 1933 gave the bound

$$\alpha_n^2 \leqq (2n-1)(n-1)^{n-1} \tag{3}$$

and 4 years later Tiberiu Popoviciu [17] sharpened this in the special case $n \equiv 1 \pmod 4$ to

$$\alpha_n^2 \leqq (n+1)\left(1+\frac{1}{n}\right)^{n-2}(n-1)^{n-1}.$$

The latter had obtained his result in terms of $(0,1)$ matrices by exploiting the properties of positive definite quadratic forms. The connection was noted in 1946 by J. Williamson [26] who showed that

$$\alpha_n = 2^{n-1}\beta_{n-1},$$

where β_n is the maximum value attained by the determinants of all $(0,1)$ $n \times n$ matrices. Presumably an $n \times n$ $(-1,1)$ matrix exists with determinant $2^{n-1}\gamma_{n-1}$ for each integer γ_{n-1}, $0 < \gamma_{n-1} < \beta_{n-1}$, but this has never been proved. During the International Symposium on Matrix Computation held in April 1961, L. Collatz asked for the maximum determinants of $(-1,0,1)$ matrices as well. A year later Ehlich and Zeller [5] noted that for each n these values will be the same as α_n.

Subsequently Ehlich [6] rederived (3) and noted that equality could hold only when

$$n = \frac{m^2+1}{2}$$

is an integer for some m. An easy computation shows that equality does hold if there is a $(-1,1)$ matrix A of order n for which $AA^T = (n-1)I_n + J_n$, where I_n is the identity matrix of order n and J is the $n \times n$ matrix whose every entry is 1. For $n = 5,13$ there are cyclic A's with maximum determinants whose first rows are given by

$$+ + + + -$$

and

$$+ + + + + - + + + - - + -,$$

where $+$ stands for $+1$ and $-$ for -1 [6]. For $n = 25$ an A with maximum determinant is known [18] but for both $n = 25$ and $n = 41$ no cyclic maximal A can exist [13].

In 1964 two papers [**6, 27**] appeared which contained the same bound for $n \equiv 2 \pmod 4$:

$$\alpha_n^2 \leqq 4(n-1)^2(n-2)^{n-2}. \tag{4}$$

Equality happens to hold in (4) if there is an A with $\det A = \alpha_n$ and

$$AA^T = \mathrm{DIAG}[B, B] \text{ where } B = (n-2)I_{n/2} + 2J_{n/2}. \tag{5}$$

Ehlich [**6**] constructed $(-1,1)$ matrices A satisfying (5) for all $n \leqq 38$ with $n \equiv 2 \pmod 4$ except $n = 22{,}34$. These were of the form

$$A = \begin{pmatrix} A_1 & A_2 \\ -A_2^T & A_1^T \end{pmatrix},$$

where A_1, A_2 are circulant matrices of order $n/2$.

In a series of papers Yang [**28, 29**] added constructions for all $n \equiv 2 \pmod 4$ up to and including 54 still excepting $n = 22$, 34 which remain the lowest undecided values when $n \equiv 2 \pmod 4$.

The problem seems more intractable in case $n \equiv 3 \pmod 4$. The bound in (3) is too large even for $n = 3$. Williamson [**26**] found that $\alpha_7 = 2^6 \cdot 9$ and $n = 11$ is the smallest integer for which the precise value of α_n is unknown. Ehlich [**7**] has determined that

$$\alpha_n^2 \leqq \frac{4 \cdot 11^6}{7^7}(n-3)^{n-7}n^7$$

for $n \equiv 3 \pmod 4$ and $n \geqq 63$.

Various functions have been used to approximate α_n for all n. Since the determinant of *any* $n \times n$ $(-1,1)$ matrix is divisible by 2^{n-1} [**8**, problem 526] and $n^{n/2}$ is attained for many known multiples of 4, one likely function is of the form

$$n^{n/2}2^{-\frac{1}{2}\psi(n)}. \tag{7}$$

Florek [**9**] has given estimates of α_n from below by estimating ψ in (7) from above. J.H.E. Cohn [**3**] used

$$n^{n/2}e^{-\phi(n)}$$

and showed [**3**] that $\phi(n) = o(n \log n)$. Lindstrom and Clements [**2**] proved that

$$\phi(n) < n \log(2 \cdot 3^{-\frac{1}{2}})$$

and J.H.E. Cohn [**4**] established that

$$\phi(n) < \begin{cases} 1 & n = p^k + 1 \\ \frac{1}{2} + \frac{1}{2}\log n & n = p^k, \end{cases}$$

where p is an odd prime. More recently another lower bound was given by Schmidt [22] in terms of (0, 1) matrices.

References

1. G. Barba, Intorno al teorema di Hadamard sui determinanti a valore massimo, Giorn. Mat. Battaglini, 71 (1933) 70–86.

2. G. F. Clements and B. Lindstrom, A sequence of (± 1)-determinants with large values, Proc. Amer. Math. Soc., 16 (1965) 548–550.

3. J. H. E. Cohn, On the value of determinants, Proc. Amer. Math. Soc., 14 (1963) 581–588.

4. ———, Determinants with elements ± 1, J. London Math. Soc., 42 (1967) 436–442.

5. H. Ehlich and K. Zeller, Binäre Matrizen, Z. Angew. Math. Phys., 42 (1962) T20–T21.

6. H. Ehlich, Determinantenabschätzungen für binäre Matrizen, Math. Zeitschr., 83 (1964) 123–132.

7. ———, Determinanten Abschätzung für binäre Matrizen mit $n \equiv 3 \bmod 4$, Math. Zeitschr., 84 (1964) 438–447.

8. D. K. Faddeev and I. S. Sominskii, Problems in Higher Algebra, Freeman, San Francisco, 1965.

9. K. Florek, On the evaluation from below of extremal determinants, Colloq. Math., 10 (1963) 111–131.

10. W. M. Frank, A bound on determinants, Proc. Amer. Math. Soc., 16 (1965) 360–363.

11. A. W. Goodman, Problem E 264, this MONTHLY, 52 (1945) 341–342.

12. J. Hadamard, Résolution d'une question relative aux déterminants, Bull. Sci. Math., 17(1893) 240–246.

13. M. Hall Jr., A survey of difference sets, Proc. Amer. Math. Soc., 7 (1956) 975–986.

14. ———, Combinatorial Theory, Blaisdell, Waltham, Mass., 1966.

15. F. Harary, Research Problem 1, Bull. Amer. Math. Soc., 68 (1962) 24.

16. G. Pall, Problem E 680, this MONTHLY, 53 (1946) 220–223.

17. T. Popoviciu, Remarques sur le maximum d'un déterminant dont tous les éléments sont non-négatifs, Soc. Sci. de Cluj., 8 (1937) 572–582.

18. D. Raghavarao, Some optimum weighing designs, Ann. Math. Statist., 30 (1959) 295–303.

19. ———, Some aspects of weighing designs, Ann. Math. Statist., 31 (1960) 878–884.

20. C. R. Rao, Factorial experiments derivable from combinatorial arrangements of arrays, J. Royal Stat. Soc. Suppl. 9 (1947) 128–139.

21. H. J. Ryser, Maximal determinants in combinatorial investigations, Canad. J. Math., 8 (1956) 245–249.

22. K. W. Schmidt, Lower bounds for maximal (0,1) determinants, Siam. J. Appl. Math. 19 (1970) 443–450.

23. J. J. Sylvester, Thoughts on inverse orthogonal matrices, simultaneous sign-successions, and tesselated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers, Phil. Mag., (4) 34 (1867) 461–475.

24. P. Turán, On a problem in the theory of determinants, Acta Math. Sinica, 5 (1955) 411–423.

25. J. Williamson, Hadamard's determinant theorem (abstract), this MONTHLY, 52 (1945) 417.

26. ———, Determinants whose elements are 0 and 1, this MONTHLY, 53 (1946) 427–434.

27. M. Wojtas, On Hadamard's inequality for the determinants of order non-divisible by 4, Colloq. Math., 12 (1964) 73–83.

28. C. H. Yang, Some designs for maximal ($+1$, -1)-determinant of order $n \equiv 2 \pmod 4$, Math. Comp., 20 (1966) 147–148.

29. ———, A construction for maximal ($+1$, -1)-matrix of order 54, Bull. Amer. Math. Soc., 72 (1966) 293.

30. ———, On designs of maximal (+ 1, −1)-matrices of order $n \equiv 2 \pmod 4$, Math. Comp., 22(1968) 174–180.

31. ———, On designs of maximal (+ 1, −1)-matrices of order $n \equiv 2 \pmod 4$ II, Math. Comp., 23 (1969) 201–205.

THE UNION OF ARITHMETIC PROGRESSIONS WITH DIFFERENCES NOT LESS THAN k

R. B. CRITTENDEN, Portland State University, and
C. L. VANDEN EYNDEN, Illinois State University

Let S be the union of n arithmetic progressions of integers, each with common difference not less than k, where $k \leqq n$. The authors conjecture that S contains all positive integers whenever it contains those not exceeding $k2^{n-k+1}$. Replacing the latter integer by $k2^{n-k+1}-1$ makes the conjecture false for any such n and k. The case $k = 1$ is known to be true.

The problem of determining the least number such that S contains all positive integers whenever it contains those not exceeding it was suggested (for $k = 3$) by Paul Erdös in a private communication. For the case $k = 1$ see **[1]**.

It is easily checked that each positive integer from 1 to $k2^{n-k+1}-1$ is a solution of either one of the $k-1$ congruences $x \equiv i \pmod{k}$, $1 \leqq i \leqq k-1$, or else one of the $n-k+1$ congruences $x \equiv 2^{j-1}k \pmod{2^jk}$, $1 \leqq j \leqq n-k+1$, but that $k2^{n-k+1}$ is not. This shows the conjecture cannot be strengthened.

Reference

1. R. B. Crittenden and C. L. Vanden Eynden, Any n arithmetic progressions covering the first 2^n integers cover all integers, Proc. Amer. Math. Soc., 24 (1970) 475–481.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

REGULARITY AS A RELAXATION OF PARACOMPACTNESS

JAMES CHEW, The University of Akron

As first introduced, the term paracompactness refers to a property of Hausdorff spaces. Normality turns out to be a necessary condition for a Hausdorff space to be paracompact. It is well known that paracompactness is equivalent to the

condition called full-normality in T_3-spaces. Since full-normality implies normality, it is possible to think of the paracompact requirement as a separation axiom stronger than normality. We experimented with a covering condition which amounts to a "localization" of the paracompact requirement and also with a condition which "localizes" the fully-normal requirement; these two conditions turn out to be equivalent to regularity in Hausdorff spaces. The purpose of this note is to prove the following theorem:

THEOREM. *The following conditions on a Hausdorff space X are equivalent*:

I. *Given an open cover $\mathscr{U}$ of X and given $a \in X$, there exists an open refinement $\mathscr{V}$ of $\mathscr{U}$ such that $\mathscr{V}$ is locally finite at a.*

II. *X is regular.*

III. *Given an open cover $\mathscr{U}$ of X and given $a \in X$, there exists an open cover $\mathscr{V}$ of X and an open set O_a containing a, such that $x \in O_a$ implies $St(x, \mathscr{V}) \subset U$ for some $U \in \mathscr{U}$.*

Condition I is obtained by weakening the paracompact condition to the requirement that, given a point *in advance*, an open cover has an open refinement that is locally finite at the specified point. Condition III is a similar relaxation of the fully-normal condition.

We shall prove the theorem by showing that I is equivalent to II for Hausdorff spaces, and that II is equivalent to III for T_1-spaces. We wish to express our gratitude to the referee for shortening the proof we originally gave to Lemma 1.

LEMMA 1. *If X is T_2 and satisfies* I, *it also satisfies* II.

Proof. Let A be a closed set and let $a \notin A$. For each $x \in A$, let O_x be an open neighborhood of x such that $a \notin \bar{O}_x$. Select an open cover $\mathscr{V}$ which refines $\{O_x : x \in A\} \cup \{X - A\}$ and an open neighborhood O_a of a which meets at most finitely many members of $\mathscr{V}$. Let

$$O_A = \cup \{V \in \mathscr{V} : A \cap V \neq \emptyset\}$$

and let $V_1, V_2, \cdots, V_m$ be the members of $\mathscr{V}$ not lying in $X - A$ which meet O_a. Then

$$O_a \cap (X - \bigcup_{i=1}^{m} \bar{V}_i)$$

is a neighborhood of a which does not meet O_A, a neighborhood of A. ∎

It is easy to show that II implies both I and III.

LEMMA 2. *If X is T_1 and satisfies* III, *it also satisfies* II.

Proof. Let $a \in X$ be given and let B be a closed set in X such that $a \notin B$. For each $x \in X$, $x \neq a$, let U_x be an open set containing x such that $a \notin U_x$. Consider the open cover $\mathscr{U} = \{X - B\} \cup \{U_x : x \in X - a\}$. Let $\mathscr{V}$ be an open cover of X

and let O_a be an open set containing a such that $x \in O_a$ implies $St(x, \mathscr{V}) \subset U$ for some $U \in \mathscr{U}$. Pick $V_a \in \mathscr{V}$ such that $a \in V_a$ and, for each $b \in B$, let $V_b \in \mathscr{V}$ be such that $b \in V_b$. Set $P_a = V_a \cap O_a$ and $P_B = \cup \{V_b: b \in B\}$. Then P_a and P_B are disjoint. For suppose $x \in P_a \cap P_B$, then $x \in P_B$ implies $x \in V_{b'}$, for some $b' \in B$. But $x \in P_a \Rightarrow x \in O_a \Rightarrow St(x, \mathscr{V}) \subset U$ for some $U \in \mathscr{U}$. If $St(x, \mathscr{V}) \subset X - B$, then $V_{b'} \subset X - B$ or $b' \in X - B$, a contradiction. If $St(x, \mathscr{V}) \subset U_y$ for $y \neq a$, then $V_a \subset U_y$ for some $y \neq a$. This contradicts the choice of U_y since we chose U_y such that $a \notin U_y$. Hence P_a and P_B are disjoint open sets containing a and B respectively. ∎

Reference

1. J. L. Kelley, General Topology, Van Nostrand, Princeton, N. J., 1961.

A SIMPLE EXAMPLE ON SOME PROPERTIES OF NORMAL RANDOM VARIABLES

JAVAD BEHBOODIAN, University of North Carolina, Chapel Hill, and Pahlavi University, Shiraz, Iran

In recent issues of this MONTHLY, several authors presented different examples on non-normality of a linear combination of normally distributed random variables [1, 2, 3]. The purpose of this note is to give a simple example of a non-normal multivariate density for demonstrating some of the properties of normal random variables. Consider

$$f(x_1, x_2, \cdots, x_n) = \sum_{k=1}^{2} p_k f_k(x_1, x_2, \cdots, x_n),$$

where $0 < p_1 < 1$, $p_1 + p_2 = 1$, and $f_k(x_1, x_2, \cdots, x_n)$ is an n-dimensional normal density function for $k = 1, 2$. The function $f(x_1, x_2, \cdots, x_n)$ is called the probability density function of a mixture of two n-dimensional normal distributions. For the sake of simplicity consider the case in which the corresponding characteristic function of $f_k(x_1, x_2, \cdots, x_n)$ is of the standard form

$$\phi_k(t_1, t_2, \cdots, t_n) = \exp\left[- \left\{ \sum_{i,j=1}^{n} \rho_{ijk} t_i t_j \right\} /2 \right],$$

where the $n \times n$ symmetric positive definite matrices $R_1 = [\rho_{ij1}]$ and $R_2 = [\rho_{ij2}]$, which are called correlation matrices, are different from each other with $|\rho_{ijk}| < 1$ for $i \neq j$ and $\rho_{ijk} = 1$ for $i = j$. It is clear that $\phi_k(t_1, t_2, \cdots, t_n)$ results from an n-dimensional normal distribution with 0 means, variances 1 and covariance matrix $R_k = [\rho_{ijk}]$. The corresponding characteristic function of $f(x_1, x_2, \cdots, x_n)$ is

$$\phi(t_1, t_2, \cdots, t_n) = \sum_{k=1}^{2} p_k \phi_k(t_1, t_2, \cdots, t_n).$$

Now suppose that $X_1, X_2, \cdots, X_n$ are n random variables with the non-normal joint density function $f(x_1, x_2, \cdots, x_n)$. Looking at the joint characteristic function $\phi(t_1, t_2, \cdots, t_n)$, the following results are obtained.

1. The random variable X_i, for $i = 1, 2, \cdots, n$, is marginally normal since its characteristic function is (see [**5**])

$$\phi(0, 0, \cdots, t_i, \cdots, 0) = \exp(-t_i^2/2).$$

However, X_i's do not have a joint n-dimensional normal distribution.

2. Let $Z = \sum_{i=1}^{n} a_i X_i$ be a linear combination of the normal random variables X_i with at least two non-zero a_i's. The characteristic function of Z, i.e.,

$$\phi_Z(t) = \phi(a_1 t, a_2 t, \cdots, a_n t) = \sum_{k=1}^{2} p_k \exp\left[\left(-\sum_{i,j=1}^{n} \rho_{ijk} a_i a_j\right) t^2/2\right]$$

shows that Z is a univariate normal if

$$\sum_{i,j=1}^{n} \rho_{ij2}\, a_i a_j \quad \text{and} \quad \sum_{i,j=1}^{n} \rho_{ij2}\, a_i a_j$$

are equal; otherwise Z has a mixture of two univariate normal distributions; and it is known that no finite mixture of two or more normal distributions can be normal [**4**]. Thus, a linear combination of normal random variables, which do not have a joint normal distribution, may or may not be normal.

3. The correlation coefficient of the normal random variables X_i, X_j is $p_1\rho_{ij1} + p_2\rho_{ij2}$. Thus, for example, if $p_1 = p_2 = \frac{1}{2}$ and $\rho_{ij1} = -\rho_{ij2} \neq 0$, we have two uncorrelated marginally normal random variables X_i and X_j which are not independent. We can easily see that the joint density of X_i and X_j is a mixture of two bivariate normal densities. This is a good example for showing the falsity of the loose statement "Two normal random variables are independent if and only if they are uncorrelated."

We now give an example which will illustrate the above results. Let $f(x_1, x_2, x_3)$ have the following correlation matrices

$$R_1 = \begin{bmatrix} 1 & 0 & -\rho \\ 0 & 1 & \rho \\ -\rho & \rho & 1 \end{bmatrix} \text{ and } R_2 = \begin{bmatrix} 1 & -\rho & \rho \\ -\rho & 1 & 0 \\ \rho & 0 & 1 \end{bmatrix}.$$

These two matrices are both positive definite if $|\rho| < \sqrt{2}/2$. The corresponding characteristic function of $f(x_1, x_2, x_3)$ is

$$\phi(t_1, t_2, t_3) = p_1 \exp[-(t_1^2 + t_2^2 + t_3^2 - 2\rho t_1 t_3 + 2\rho t_2 t_3)/2] + p_2 \exp[-(t_1^2 + t_2^2 + t_3^2 - 2\rho t_1 t_2 + 2\rho t_1 t_3)/2].$$

We observe that, for the normal variables X_1, X_2, X_3, the linear combination $Z_1 = X_1 + X_2 + X_3$ is normal while the linear combination $Z_2 = X_1 - X_2 + X_3$ is not normal. We also notice that, for $p_1 = p_2 = \frac{1}{2}$, the normal variables X_1 and X_3 are uncorrelated while they are not independent.

It is clear that similar results are obtained even if we use a finite mixture of more than two n-dimensional distributions.

References

1. G. E. Albert and R. L. Tittle, Non-normality of linear combinations of normally distributed random variables, this MONTHLY, 74 (1967) 583–585.
2. B. K. Kale, Normality of linear combinations of non-normal random variables, this MONTHLY, 77 (1970) 992–995.
3. Lloyd Rosenberg, Non-normality of linear combinations of normally distributed random variables, this MONTHLY, 72 (1965) 888–890.
4. H. Teicher, On the mixture of distributions, Ann. Math. Stat., 31 (1960) 55–73.
5. S. S. Wilks, Mathematical Statistics, Wiley, New York, 1962.

AN ALTERNATIVE TO THE INTEGRAL TEST FOR INFINITE SERIES

G. J. PORTER, University of Pennsylvania

Infinite series are usually studied in a calculus course following the development of the integral. One reason for this placement is the desire to have the integral test available. An earlier study of infinite series might be desired to complement the study of sequences or to study Taylor series as an immediate application of the derivative. In these cases an alternative to the integral test is needed.

One alternative is the Cauchy Condensation Test; this method seems to be well known in Europe and Latin America, but not in the United States. Many calculus teachers are aware that this test (perhaps not by this name) may be used to prove that the series $\Sigma\ 1/n$ diverges. I suspect a smaller number are aware that it may be used for all the series which are usually studied by the integral test. It is the point of this note to recall the test and give several examples of its use.

THEOREM (Cauchy Condensation Test). *Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms such that $a_{n+1} \leqq a_n$ for all n. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if the condensed series $\sum_{j=1}^{\infty} 2^j a_{2^j}$ converges.*

Proof. Since

$$2^{j-1}a_{2^j} \leqq a_{2^{j-1}+1} + a_{2^{j-1}+2} + \cdots + a_{2^j} \leqq 2^{j-1}a_{2^{j-1}}$$

we have

$$\sum_{j=1}^{\infty} 2^{j-1}a_{2^j} \leqq \sum_{n=2}^{\infty} a_n \leqq \sum_{j=1}^{\infty} 2^{j-1}a_{2^{j-1}} \leqq 2 \sum_{j=1}^{\infty} 2^{j-1}a_{2^j}$$

and the theorem follows.

Example 1: $\Sigma_{n=1}^{\infty} 1/n^{\alpha}$. The condensed series is

$$\sum_j \frac{2^j}{(2^j)^{\alpha}} = \sum_j \frac{1}{(2^j)^{\alpha-1}} = \sum_j \frac{1}{(2^{\alpha-1})^j}.$$

This is a geometric series and converges if and only if $2^{1-\alpha} < 1$, i.e., $\alpha > 1$. Thus the given series converges if and only if $\alpha > 1$.

Example 2: $\Sigma_{n=2}^{\infty} 1/n(\log n)^{\alpha}$. The condensed series is

$$\Sigma_j \frac{2^j}{2^j(\log(2^j))^{\alpha}} = \Sigma_j \frac{1}{(\log(2^j))^{\alpha}} = \Sigma_j \frac{1}{(j\log 2)^{\alpha}} = \frac{1}{(\log 2)^{\alpha}} \Sigma_j \frac{1}{j^{\alpha}}$$

which converges if and only if $\alpha > 1$ by Example 1.

Example 3. $\Sigma_{n=2}^{\infty} 1/n\log n(\log(\log n))^{\alpha}$. The condensed series is

$$\Sigma_j \frac{2^j}{2^j\log(2^j)(\log(\log 2^j))^{\alpha}} = \Sigma_j \frac{1}{j(\log 2)(\log(j\log 2))^{\alpha}}$$

$$= \frac{1}{\log 2} \Sigma \frac{1}{j(\log j + \log 2)^{\alpha}}$$

which converges if and only if $\alpha > 1$ by comparison with Example 2.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

MATHEMATICS COURSES IN 1984

J. BARKLEY ROSSER, University of Wisconsin

To answer the question, "What Undergraduate Courses Will be Taught in 1984?" one proceeds through four assessments: (a) what will be taught if past trends simply continue? (b) what ought to be taught? (c) what steps can be taken to convert some of (a) to (b)? (d) how many of the steps proposed under (c) will be undertaken, and what success will they have?

2. If for "omitting proofs" we read "omitting demonstrations" then this method of teaching is common in all the sciences. Most students in physics and chemistry, for example, seldom perform the basic experiments which support modern theory. The analogy to mathematics is not precise, since to the practising mathematician the method of proof is frequently as valuable as the final result. But the non-mathematical scientist generally must know only when to apply certain mathematical techniques. Here an intuitive understanding of the rationale without precise demonstrations will go a long way towards satisfying his needs.

3. There are several good examples of such "boot strap" operations of which the University of Pennsylvania is one. Starting with one experimental computer-assisted section serving approximately 10% of the freshmen, it has become possible to offer integrated calculus-with-computer courses to all who desire it, and these are in the majority. The main difficulty is not in professors learning to program — enough will volunteer — but in getting enough teaching assistants to correct the students' mistakes.

4. Some view the computer as the greatest single emerging threat to civil liberties. There is little possibility of turning computers off. Intelligent regulation of their use will require a body of informed citizens who understand what they can do. Some instruction in the use of computers is probably therefore in order even for pre-law students. For a well-documented account of present dangers see the work of Professor Arthur Miller of the University of Michigan Law School, "The Assault on Privacy — Computers, Data Banks, and Dossiers", University of Michigan Press, Ann Arbor, 1971.

PROBLEMS AND SOLUTIONS

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before September 30, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2361. *Proposed by Richard Johnsonbaugh, Morehouse College*

Prove that the following series converge conditionally:

$$\sum_{n=1}^{\infty}(-1)^n(n^{1/n}-1) \text{ and } \sum_{n=1}^{\infty}(-1)^n[e-(1+1/n)^n].$$

E 2362.* *Proposed by C. H. Kimberling, University of Evansville*

Suppose that in some probability space, E_1, $E_2,\cdots$ are events with common probability p. Let $m \geqq 2$ be a fixed integer. Prove or disprove that

$$p^m \leqq \sup\{P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m})\},$$

where the supremum is taken over all m-tuples $(i_1, i_2, \cdots, i_m)$ of distinct natural numbers.

E 2363. *Proposed by Hüseyin Demir, Middle East Technical University, Ankara, Turkey*

Characterize pairs of spherical triangles ABC and $A'B'C'$ for which $A'=a$, $B'=b$, $C'=c$, $A=a'$, $B=b'$, $C=c'$.

E 2364. *Proposed by G. J. Michaelides, University of South Florida.*

Suppose that r is a positive integer and that $(i_1, i_2,\cdots, i_n)$ is a partition of r into nonnegative integers. Show that if p is a prime factor of n which is relatively prime to r, then the number of (distinct) permutations of $(i_1, i_2, \cdots, i_n)$ is divisible by p.

E 2365. *Proposed (part I) by Erwin Just and Kenneth Fogarty, Bronx Community College, and (part II) by J. B. Wilker, University of Toronto*

I. Let S be a finite set of points in the plane in which no three points are collinear and not all points are concyclic. Define a *common* point of S to be a point which lies on some circle which passes through precisely two other points of S. Must each point of S be a common point?

II. Let S be a set of four or more points lying on a sphere but not on a circle. Prove that each point of S is on some circle containing precisely two of the other points of S.

E 2366.* *Proposed by B. P. Gill, Demarest, N.J.*

Let V be the set of vertices of a regular $2n$-gon and let A^* and B^* be convex n-gons whose vertices are subsets A and B of V. If the set of lengths of all chords with both ends in A (with each chord length being counted according to its multiplicity) is identical to the like set for B, then is B^* necessarily congruent to either A^* or $(V\setminus A)^*$?

SOLUTIONS OF ELEMENTARY PROBLEMS

Wilsonian Products in a Group

E 2303 [1971, 674]. *Proposed by Charles Lindner, Auburn University*

Let G be a finite group of odd order. Then the set of products of all elements of G, taken in any order, is in the commutator subgroup. [As a corollary we have the well-known result that G abelian implies $\Sigma_{x \in G}\, x =$ identity. Query: Does the set of all such products exhaust the commutator group?—Ed.]

I. *Solution by G. A. Heuer, Concordia College.* Let $G = \{g_1, g_2, \cdots, g_n\}$, let N be the commutator subgroup of G, and suppose that the order of N is m. Then $(g_1 g_2 \cdots g_n)N = (g_1 N)(g_2 N) \cdots (g_n N) = X^m$, where X is the product of all elements of G/N; since G/N is an abelian group of odd order, $X = N$, the identity of G/N. Thus $g_1 g_2 \cdots g_n \in N$.

II. *Comment by Solomon Golomb, University of Southern California.* While the problem posed is indeed elementary, the "Editor's Query" is anything but elementary. I have been interested in this question (as applied to *all* finite groups) since it first occurred to me in 1951; finally around 1967, having despaired of solving it myself, I submitted it as a Research Problem to the Bulletin of the AMS, where it languished for several years, finally appearing there in vol. 76, no. 5, September 1970, as problem 8 on pp. 973–974, entitled "Wilsonian products in groups." I have received no solutions to date.

Also solved (first part only) by Ram Avtar (India), Anders Bager (Denmark), Michael Barr (Denmark), S. Baskaran (India), California Polytechnic Solutions Group, Fred Clare, John Coolidge, Harold Donnelly, S. F. Ebey, Daniel Farkas, Bruce Ferrero, Zbigniew Fiedorowicz, S. W. Golomb, M. G. Greening (Australia), J. W. Grossman, Elgin Johnston, Geoffrey Kandall, David Kelly, Yuriko Kojima, Harry Lass, C. B. A. Peck, Ernest Propes, Simeon Reich (Israel), Azriel Rosenfeld, Daniel Shapiro, Stephen Spindler, Glenn Stevens, John Stout, D. P. Sumner, E. T. Wong, and the proposer.

Editor's Comment. It is widely known that if G is abelian, then the product of the elements of G is the identity, except in the case that there is a unique element $x \in G$ of order two; in this case the product is x. Applying this to the multiplicative group of numbers mod p, where p is a prime, we have that $(p-1)! \equiv -1 \pmod p$, which is Wilson's Theorem. This is the origin of the term "Wilsonian product".

S. Baskaran and C. B. A. Peck refer to A. R. Rhemtulla, *On a problem of L. Fuchs,* Studia Scientarum Mathematicarum Hungarica, 4 (1969), 195–200. If G has order n and if S is the set of all products of n distinct elements (i. e., the set of all Wilsonian products), then it is not hard to show that S is contained in a single coset of the commutator subgroup. Rhemtulla shows that S is equal to that coset in the case that G is solvable. In this paper, mention is made of work by J. Dénes which shows that the same result holds when every member of the commutator subgroup is actually a commutator and when in addition, G has at most $\frac{1}{2}n$ elements of order two.

Cardinality of Intersecting Cosets

E 2304 [1971, 674]. *Proposed by J. C. Owings, Jr., University of Maryland*

Let G be a finite group, H a subgroup of G. Show that, given any left coset L of H, there exists an integer k such that, for any right coset R of H, $L \cap R$ is either empty or has cardinality k.

Solution by the Bennett College Team. We prove the more general assertion: *Let G be a group (finite or not) and let H be a subgroup of G. If L is any left coset of H, then there exists a cardinal number k such that if R is any right coset of H, then $L \cap R$ is either empty or has cardinality k.* To prove this, suppose that $L = aH$. The right cosets that meet L are precisely those of the form Hah where $h \in H$; one of them is Ha. The proposition will follow if we can show for each $h \in H$ there exists a bijection from $aH \cap Ha$ to $aH \cap Hah$. Given such an h, let $F(x) = xh$ for all $x \in aH \cap Ha$. Clearly F is an injection from $aH \cap Ha$ to $aH \cap Hah$. We now show that F is a surjection: If $y \in aH \cap Hah$, then $y = ah_1 = h_2ah$ for some h_1, $h_2 \in H$. Let $x = h_2a$; then $x \in Ha$ and $x = h_2a = ah_1h^{-1} \in aH$. But $F(x) = xh = h_2ah = y$, and the assertion is proved.

Also solved by twenty-four other readers and the proposer.
Several other solvers also dispensed with the finiteness condition.

Stern Rediscovered

E 2305 [1971, 674]. *Proposed by M. D. Hendy, University of New England, Australia*

In the system of reduced residues modulo p, where p is a prime, for each $e \mid p-1$, there are $\phi(e)$ elements of order e. Prove that the sum of these $\phi(e)$ elements $\equiv \mu(e)$ $(\operatorname{mod} p)$, where $\phi(e)$ is the Euler totient function and $\mu(e)$ is the Möbius function.

I. *Solution by Leonard Carlitz, Duke University.* If e divides $p-1$, let $S(e)$ denote the sum of the $\phi(e)$ numbers belonging to the exponent e $(\operatorname{mod} p)$ and let $T(e)$ denote the sum of the e numbers satisfying $x^e \equiv 1$ (all congruences being $\operatorname{mod} p$). Obviously $T(e) \equiv \sum_{d|e} S(d)$ so that (as in the Möbius inversion formula), $S(e) \equiv \sum_{d|e} T(d)\mu(e/d)$. But $T(e)$ is the sum of all solutions of $P(x) = x^e - 1 \equiv 0$ in the integers $\operatorname{mod} p$; this is the negative of the coefficient of x^{e-1} in $P(x)$ so that $T(e) \equiv 0$, if $e > 1$, and $T(1) \equiv 1$. Substitution shows that $S(e) \equiv \mu(e)$.

If we let $S_k(e)$, denote the sum of the kth powers of the $\phi(e)$ numbers belonging to the exponent $e (\operatorname{mod} p)$ and define $T_k(e)$ analogously, then we can show by a similar argument that $S_k(e) \equiv \sum d\mu(e/d)$, the sum being taken over all divisors d of (e,k). (Cf. E. Landau, *Vorlesungen über Zahlentheorie* I, Satz 220, p. 188.)

II. *Solution by Solomon Golomb, University of Southern California.* Let $\Phi_n(x)$ denote the cyclotomic polynomial of order n. (That is, $\Phi_n(x) = \Pi(x - \zeta)$,

where ζ runs through the primitive nth roots of unity.) It is known that $\Phi_n(x)$ has integer coefficients and if $s = \phi(n)$, then $\Phi_n(x) = x^s - \mu(n)x^{s-1} \pm \cdots$; that is, the degree of $\Phi_n(x)$ is $\phi(n)$, and $\mu(n) = \sum \zeta$, where ζ runs through the primitive nth roots of unity. (This is the "algebraist's definition" of the Möbius function.)

For any prime p, the natural modulo p mapping from $Z[x]$ to $Z_p[x]$ is a ring homomorphism, and $\Phi_n(x)$ with coefficients reduced mod p is still the polynomial for the primitive nth roots of unity over Z_p, provided $(n, p) = 1$. (We remark that it is possible to have coefficients other than 0, 1, or -1 in the cyclotomic polynomial, despite occasional printed statements to the contrary.) In particular, if n divides $p-1$, then surely $(n, p) = 1$; moreover, the $\phi(n)$ primitive nth roots of unity are elements of Z_p itself, and their sum must equal the negative of the second coefficient of $\Phi_n(x)$ (mod p), which is $\mu(n)$.

Also solved by twenty-seven other readers and the proposer.

Editor's comment. According to L. E. Dickson, *History of the Theory of Numbers*, Vol. I, Chelsea, New York, 1952 (p. 184), this result was known to M. A. Stern (Jour. für Math. 6 (1830), p. 258). Ney Borba comments that the case $e = p - 1$ was proved by Gauss (*Disq. Arith.* I, Article 81).

A Singular Problem

E2306 [1971, 674]. *Proposed by Anon, Erehwon-upon-Wabash*

Let A be an $n \times n$ matrix, $\boldsymbol{u}$ a $1 \times n$ row vector, $\boldsymbol{v}$ an $n \times 1$ column vector, and

$$B = \begin{bmatrix} A & -A\boldsymbol{v} \\ -\boldsymbol{u}A & \boldsymbol{u}A\boldsymbol{v} \end{bmatrix}.$$

(a) Prove that 0 is a characteristic root of B.

(b) Suppose $\det A = 0$. Show that t^2 divides the characteristic polynomial $\det(tI - B)$ of B.

(c) Discuss the converse of (b).

Solution by W. G. Leavitt, University of Nebraska. To show (a), simply note that $(\boldsymbol{u}\,|\,1)B = 0$. To show (b), suppose that $\det A = 0$ so that $\boldsymbol{x}A = 0$ for some row vector $\boldsymbol{x} \neq 0$. Then $(\boldsymbol{x}\,|\,0)B = 0$; but obviously $(\boldsymbol{x}\,|\,0)$ and $(\boldsymbol{u}\,|\,1)$ are independent, so that t^2 divides the characteristic polynomial $P(t) = \det(tI - B)$ of B. For (c), define P as follows:

$$P = \begin{bmatrix} I & 0 \\ \boldsymbol{u} & 1 \end{bmatrix}.$$

Then B has the same characteristic polynomial as PBP^{-1}; it is easily seen that

$$PBP^{-1} = \begin{bmatrix} A(I + \boldsymbol{v}\boldsymbol{u}) & -A\boldsymbol{v} \\ 0 & 0 \end{bmatrix},$$

so that $P(t) = \det(tI - B) = t\det(tI - A(I + \boldsymbol{vu}))$. Now t^2 divides $P(t)$ if and only if t divides $\det(tI - A(I + \boldsymbol{vu}))$, which means that either A or $(I + \boldsymbol{vu})$ is singular. Since $\det(I + \boldsymbol{vu}) = 1 + \boldsymbol{uv}$, it follows that t^2 divides $P(t)$ if and only if $\det A = 0$ or $\boldsymbol{uv} = -1$.

Also solved by thirty-three other readers and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before September 30, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

Problem 5837 [1972, 94] is withdrawn. Inexplicably Problem 5797 reappears with a new number.

5860*. *Proposed by L.-S. Hahn, University of New Mexico*

Let $f(z)$ be a measurable function in the plane, assumed integrable on all circles with radius 1. Suppose $f(z)$ has the property that its value at an arbitrary point in the plane is the average of its values on the circle of radius 1 centered at that point, viz.

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} f(z + e^{i\theta})\,d\theta \quad \text{for all } z \in C.$$

Is the function $f(z)$ necessarily continuous?

5861*. *Proposed by Michael Slater, University of Bristol, England*

Let F be an ordered field.

(a) If $p \in F[x]$; $a, b \in F$, $a < b$, and $p'(x) > 0$ for $a \leqq x \leqq b$, does it follow that $p(a) < p(b)$?

(b) If Rolle's theorem holds in F, does it follow that F is real-closed?

5862. *Proposed by R. C. Wagner, Fairleigh Dickinson University*

A submodule N of the R-module M is said to be *pure* if for every $r \in R$, $rN = N \cap rM$. Prove that if R is a commutative Noetherian ring with unit and M is a finitely generated R-module for which every submodule is pure, then every submodule is a direct summand of M.

5863. *Proposed by P. R. Chernoff, University of California, Berkeley.*

Let D be an integral domain with infinitely many elements. Assume that every

non-unit in D has an irreducible factor. Prove that D has infinitely many irreducibles or infinitely many units.

5864. *Proposed by G. E. Andrews, Pennsylvania State University*

Let P_n denote the set of partitions of n into positive integers. For each $\pi \in P_n$, let $d(\pi)$ denote the number of different parts of π, and let $\#(\pi)$ denote the total number of parts of π. Prove that for $n \geqq 1$,

$$\sum_{\pi \in P_n} (-1)^{\#(\pi)} 2^{d(\pi)} = \begin{cases} 2(-1)^n & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}$$

5865. *Proposed by G. E. Andrews, Pennsylvania State University*

Let Q_n denote the set of partitions of n into distinct non-negative parts with an even number as the smallest part. Let $q_e(n)$ (resp. $q_o(n)$) denote the number of elements of Q_n that have an even number (resp. odd number) of even parts. Prove that

$$q_o(n) - q_e(n) = \begin{cases} 1 & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}$$

SOLUTIONS OF ADVANCED PROBLEMS

Finite Groups with Related Generators

5788 [1971, 305]. *Proposed by N. S. Mendelsohn, University of Manitoba*

Let G be a group with presentation $G = (a, b\colon a = (ba)^r b,\ b = (ab)^s a)$. Show that G is finite for all choices of the positive integers r and s, and that either G is cyclic or G has a cyclic subgroup of index 2.

Solution by Roy Olson, University of Washington. From the relations $a = (ba)^r b$ and $b = (ab)^s a$, obtain $ab = (ba)^{r+s+1}$and $ba = (ab)^{r+s+1}$. Then $[(ab)^{r+s+1}]^{(r+s+1)} = ab$. Let H be the subgroup of G generated by ab. H is cyclic, finite, and contains ab, ba, $ab^{-1} = (ba)^r$, $a^2 = ab^{-1}ba$, $a^{-1}b = a^{-2}ab$, b^2. If the coset $aH \neq H$, then a^2H, baH, and $b^{-1}aH$ are all equal to H. Further, $bH = b(b^{-1}aH) = aH$. If $a \in H$, then $b = a^{-1}ab \in H = G$. Therefore H has index $\leqq 2$, and G is finite since H is.

Also solved by W. O. Alltop, James Alonso, G. W. Fehlhaber, L. T. Gardner, R. W. Gatterdam, J. D. Gillam, M. G. Greening (Australia), Fletcher Gross, C. V. Heuer & G. A. Heuer, D. A. Leonard, L. E. Shader, M. J. Wicks (Singapore), and Mark Yu.

Gross offers a more detailed description of the group G without restricting r and s to be positive:

1. If $r + s \neq 0$ and $(r+1)^2 + (s+1)^2 \neq 0$, then G is finite of order $|r+s| \cdot (r+1, s+1)$. G is

abelian if and only if $(r+1, s+1) = 1$. G is cyclic if and only if $(r+1, s+1) = 1$ and either $r \equiv s \pmod 2$ or $r \equiv s \equiv 0 \pmod 4$.

2. If $r = s = -1$, then G is the infinite dihedral group.

3. If $r + s = 0$ and $r \equiv 0 \pmod 2$, then G is infinite cyclic.

4. If $r + s = 0$ and $r \not\equiv 0 \pmod 2$, then G is the direct product of an infinite cyclic group and a group of order 2.

Measurable Sets which Contain No Rectangles

5789 [1971, 410]. *Proposed by P. C. Shields, Menlo Park, California*

If A and B are measurable subsets of the unit interval, then $A \times B$ is called a rectangle. Find a measurable subset of the unit square which is not a countable union of rectangles, except for a set of measure zero.

Solution by R. C. Weger, South Dakota School of Mines and Technology. Let C be a Cantor-like set which is a subset of the unit interval with positive measure, that is, C is closed, has void interior and is of positive measure. Let

$$S = \{(x, y) \mid x - y \in C, \quad 0 \leqq x \leqq 1, \quad 0 \leqq y \leqq 1\}.$$

Then if $A \times B \subseteq S$ it follows that either A or B has zero measure. For $A \times B \subseteq S$ implies that $A - B = \{x - y \mid x \in A,\ y \in B\} \subseteq C$. And if A and B were both of positive measure then $A - B$ would have nonvoid interior.

The result follows as S has positive measure.

Also solved by R. O. Davies (England), Richard Gisselquist, Joel Levy, Jan Mycielski, J. C. Oxtoby, and B. L. Schwartz.

Notes. (1) Davies generalizes by showing that if $(X, \mathbf{X}, \mu)$ and $(Y, \mathbf{Y}, \nu)$ are any two non-atomic probability measure spaces, then there exists a set $E \in \mathbf{X} \times \mathbf{Y}$ with $(\mu \times \nu)(E) > 0$, such that $(\mu \times \nu)(R \setminus E) > 0$ for every rectangle R of positive $(\mu \times \nu)$-measure.

(2) Schwartz and Oxtoby refer us for a solution to the note of Darst and Goffman, *A Borel set which contains no rectangles*, this MONTHLY, 77 (1970) 728.

(3) In addition, Oxtoby notes that the solution may also be found in a paper by himself and P. Erdös, Trans. A. M. S., 79 (1955) 91–102, Theorem 1. In this paper it is also proved that subsets of the square which are equivalent modulo nullsets to some countable union of measurable rectangles constitute only a set of first category in the space of measurable subsets of the square.

Skew-Symmetric Second Order Directional Differential

5791[1971, 410]. *Proposed by M. Z. Nashed, Georgia Institute of Technology*

For $f: R^3 \to R$, $x_0 \in R^3$, and nonzero $h_1, h_2 \in R^3$, the first and second directional derivatives are defined by

$$\delta f(x_0; h_1) = \lim_{t \to 0} \frac{1}{t}\{f(x_0 + th_1) - f(x_0)\},$$

and

$$\delta^2 f(x_0; h_1, h_2) = \lim_{t\to 0} \frac{1}{t}\{\delta f(x_0 + th_2; h_1) - \delta f(x_0; h_1)\},$$

whenever these limits exist.

Construct a function $f: R^3 \to R$ for which $\delta^2 f(x_0; h_1, h_2)$ is a skew-symmetric nonzero bilinear form at some $x_0 \in R^3$ (i.e., $\delta^2 f(x_0; h_1, h_2) = -\delta^2 f(x_0; h_2, h_1)$ for all $h_1, h_2 \in R^3$, and $\delta^2 f(x_0; h_1, h_2)$ is linear in h_1 and h_2 separately), or show that such a function does not exist.

Solution by R. L. Van de Wetering, San Diego State College. Let $f: R^3 \to R$ be given by

$$f(x,y,z) = xy\frac{x^2 - y^2}{x^2 + y^2} + xz\frac{x^2 - z^2}{x^2 + z^2} + yz\frac{y^2 - z^2}{y^2 + z^2},$$

or

$$f(0,0,z) = f(0,y,0) = f(x,0,0) = 0.$$

Now

$$\lim_{x\to 0}\frac{f(x,y,z) - f(0,y,z)}{x} = f_x(0,y,z) = -y - z.$$

Similarly we get $f_y(x,0,z) = x - z$ and $f_z(x,y,0) = x + y$. From this it follows that $f_{xy}(0,0,0) = -1$, $f_{yx}(0,0,0) = 1$, $f_{xz}(0,0,0) = -1$, $f_{zx}(0,0,0) = 1$, $f_{zy}(0,0,0) = 1$, $f_{yz}(0,0,0) = -1$. We also have, after calculating $f_x(x,y,z)$, that

$$f_{xx}(0,0,0) = \lim_{x\to 0}\frac{f_x(x,0,0) - f_x(0,0,0)}{x} = 0.$$

Similarly, $f_{yy}(0,0,0) = f_{zz}(0,0,0) = 0$.

Finally $\delta^2 f(0; h, k) = h_2k_1 + h_3k_1 - h_1k_2 + h_3k_2 - h_1k_3 - h_2k_3$ which is a skew-symmetric bilinear form given by the matrix

$$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}.$$

Fundamental Group of Non-orientable Manifolds

5792 [1971, 411]. *Proposed by W. S. Massey, Yale University*

It is well known that given any finitely presented group G and any integer $n \geqq 4$, there exists a compact, orientable n-manifold M^n such that its fundamental group, $\pi_1(M^n)$, is isomorphic to G. Is an analogous theorem true for non-orientable

manifolds? An obvious necessary condition is that G have a subgroup of index 2, since the set of orientation preserving path classes in a non-orientable manifold constitutes a subgroup of its fundamental group which is of index 2.

Solution by the proposer. The following theorem is the desired analogue:

THEOREM. *Let G be a finitely presented group, H a subgroup of G of index* 2, *and n an integer* $\geqq 4$. *Then there exists a compact non-orientable n-manifold M and an isomorphism ϕ of $\pi_1(M)$ onto G such that ϕ maps the subgroup of orientation preserving paths onto H.*

Sketch of Proof: The proof is somewhat similar to the proof of the analogous result for orientable manifolds (cf. *Algebraic Topology*: *An Introduction*, by W. S. Massey, pp. 143–144). Let Z_2 denote a cyclic group of order 2 and $w_1: G \to Z_2$ the unique homomorphism that has kernel H; we shall arrange the construction so that w_1 will be the first Stiefel-Whitney class of M. Let $g_1, \cdots, g_n$ be generators for G and $r_1, \cdots, r_m$ relations for G; each r_i is an element of the free group F generated by $g_1, \cdots, g_n$. Let $h: F \to G$ denote the natural homomorphism; the kernel of F is the normal subgroup generated by the relations r_i.

Corresponding to each generator g_i, we shall choose a compact n-manifold M_i as follows: if $w_1 h(g_i) = 0$, then $M_i = S^1 \times S^{n-1}$, while if $w_1 h(g_i) \neq 0$, $M_i =$ an n-dimensional Klein bottle (i.e., a non-orientable $(n-1)$-sphere bundle over S^1). In either case, the fundamental group of M_i is infinite cyclic. Let M' denote the connected sum of the M_i; then $\pi_1(M') = F$, and the first Stiefel-Whitney class of M' "realizes" the homomorphism $w_1 h: F \to Z_2$. Corresponding to each relation $r_i \in F$' choose a smooth imbedding f_i: $S^1 \to M'$ which represents the corresponding element of $\pi_1(M')$. Since $h(r_i) = 0$, it follows that $w_1 h(r_i) = 0$; hence the closed path f_i is orientation preserving and the normal bundle of the imbedding f_i is trivial. Thus we can do surgery on each of the imbeddings f_i (see e.g., C. T. C. Wall, *Surgery on Compact Manifolds*, New York, 1970) and obtain a manifold M with the desired fundamental group; the details are similar to the proof of the analogous theorem in the orientable case.

We note a corollary of this result. Given any finitely presented group H and integer $n \geqq 4$, there exists a compact, orientable n-manifold whose fundamental group is isomorphic to H, and which admits a fixed point free orientation reversing smooth involution.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 7

CONTENTS

(*Continued on inside cover*)

AUGUST–SEPTEMBER **1972**

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

SOME MATHEMATICAL MODELS OF POPULATION GENETICS

SAMUEL KARLIN, The Weizmann Institute of Science, Israel, and

Stanford University, California

Introduction. Theoretical population genetics and mathematical genetics is the study of temporal and spatial changes of frequencies of types (e.g., genes, genotypes, gametes, etc.) in populations subject to various ecological and genetic influences. Two general opposite tendencies operate on natural population: (i) propensity for adaptability and persistence of specific types favorable to a given environment, and (ii) necessity for populations to maintain potential for variation to cope with situations of changing environments.

The use of mathematics in studying genetic systems is as old as the subject of genetics itself. From the rediscovery of Mendel's work at the beginning of this century it did not take long for the Hardy-Weinberg law (1908)* on the constancy of gene frequency over time to be enunciated. Between 1915 and 1950 mathematical genetics was pioneered and dominated by the names of R. A. Fisher, S. Wright, and J. B. S. Haldane.

The challenge to understand the role of such genetic and ecological factors as mutation and migration rates, the varied manifestations of natural selection, the effects of population behavior and mating patterns, the relevance of recombination, etc., motivated these men to formulate a vast hierarchy of mathematical models describing many facets of population genetic phenomena. Relatively few of these models have as yet yielded to complete analysis.

Haldane, in his famous series of papers in the Proceedings of the Cambridge Philosophical Society in the 1920's, set forth a variety of simple mathematical analyses concerned with the way natural selection might be supposed to act. In particular, he indicated how evolutionary forces such as viability selection, mutation, migration, and sex-linkage could be quantified and brought into these models.

Samuel Karlin received his Princeton Ph.D. under S. Bochner. He has held positions at Cal Tech, Princeton, Stanford, and the Weizmann Institute of Science. At various times he held the Proctor Fellowship, Bateman Fellowship, Wald Memorial Lectureship, Guggenheim Fellowship, and National Science Senior Fellowship. He is a Fellow of the International Statistical Institute, the Institute of Mathematical Statistics, an elected member of the U.S. National Academy of Sciences, and the American Academy of Arts and Sciences.

Professor Karlin has been most productive in a variety of fields. He has supervised 35 Doctoral students, many now recognized scientists, has written over 125 research papers and the following books: *Studies in the Mathematical Theory of Inventory and Production* (with K. Arrow and H. Scarf, Stanford Univ. Press, 1958); *Mathematical Methods and Theory in Games, Programming, Economics,* Volume I: *Matrix Games, Programming and Mathematical Economics,* (Addison-Wesley, 1959); *Mathematical Methods and Theory in Games, Programming, Economics,* Volume II: *The Theory of Infinite Games* (Addison-Wesley, 1959); *A First Course in Stochastic Processes* (Academic Press, 1966); *Tchebycheff Systems: With Applications in Analysis and Statistics,* (with W. J. Studden, Interscience, 1966); and *Total Positivity, Volume I,* (Stanford Univ. Press, May 1968). *Editor.*

*This is the G. H. Hardy of mathematical fame.

Fisher and Wright were also involved in the elaboration of these theories. Wright further established that in small populations, evolutionary theory should take account of the sampling effects involved in producing one generation from the previous. He called this effect "random drift". This aspect of population genetics has had significant mathematical consequences especially in stimulating Feller's investigations into boundary theory of diffusion processes on the line.

Again it was Wright and Fisher who pioneered the theory of systems of mating between relatives, such as used by animal and plant breeders. The result was the theory of inbreeding which entails intriguing algebraic and analytic structures much of which is not well understood. Statistical theory probably owes its origin to R. A. Fisher's attempts to design and analyze experiments whose purposes were most often to solve problems in genetics.

The objective of this paper is to acquaint the mathematics student with several classical mathematical genetic models. Attention is mainly given to the formulation of the models accompanied by brief analyses and appropriate references. Some interpretations and implications of the results with reference to evolutionary theory are appended. On occasion relevant unsettled mathematical problems are noted.

It should be underscored that the array of models to be discussed is a very slight representation of the vast number formulated and partly dealt with by geneticists over the past half century and very recently by some mathematicians. We have attempted to highlight several important genetic factors and concepts by presenting models involving different mating patterns, selective forces, migration and mutation pressures, the recombination mechanism, etc. Many types of mathematical genetic models have been omitted in this expository article for lack of space. For example, we avoided entirely the enticing and important excursion into stochastic genetic models. (The interested reader can consult Crow and Kimura [7], Chapters 10–12, for an introduction to this part of mathematical genetics, and references cited therein.) Models based on statistical genetics have also been left out. The general theory of inbreeding systems is given scant attention (see Karlin [**16**] and [**17**] for a fuller treatment of this subject). The extensive and important literature of genetic traits determined by several loci is only briefly touched on in Section 8. (For a review on this current very active topic, consult Kojima and Lewontin [**27**], see also Karlin and Feldman [**19**], and Karlin [**20**].)

In closing the introduction, we indicate the organization of the paper. Section I reviews succinctly some of the basic terminology and relevant genetic mechanism. Section II covers a few basic random mating models exhibiting selection balance. Sections III and IV highlight two important situations of non-random mating. Section III is specially devoted to an exposition of some models involving positive assortative mating while Section IV exposes the phenomena of incompatibility mechanisms in mating patterns. These include cases of self-sterility and sex determination. Section V presents briefly the classical model of mutation selection balance

for two alleles (alternative gene forms). Section VI is concerned with the very useful method and concept of identity by descent. Section VII discusses some models of the evolution of a population with an infinite number of possible types. Section VIII introduces the simplest two locus selection model.

I. PERTINENT GENETIC PRELIMINARIES

It is unfortunate but necessary to learn a minimum of the terminology and mechanisms of population genetic systems. **Chromosomes**—usually found in the nucleus—mostly govern the inheritable characteristics of an organism. Chromosomes may occur singly (the **haploid case**) as in some fungi, in pairs (the **diploid case**), as in mammals, or in larger groups (**triploid**, **tetraploid**, in general **polyploid**) as in many plants. The associated pairs, triplets, etc., of chromosomes are called **homologous.** **Locus** is the position at which a gene (a sort of unit of the chromosome) occurs on a chromosome. **Alleles** are alternate gene forms at a given locus. **Genotypes** are the various possible combinations of alleles at corresponding loci on homologous chromosomes. In the diploid case if the alleles are A and a, the genotypes are AA, Aa, and aa.

The populations to be considered here, unless specified otherwise, contain diploid individuals. We concentrate our attention, for the most part, on characters determined by one or two loci, on a given pair of chromosomes. We usually assume that two alternative genes (alleles) may occur at each locus. Consider the case of two loci, where the alleles A and a are possible at the first locus and alleles B and b at the second locus. A typical one of the ten possible genotypes (see listing immediately below) could be written AB/ab. The symbol AB/ab signifies that AB sit on one chromosome A at the first locus, B at the second locus and ab are situated on the second chromosome. The ten genotypes are explicitly

$$\frac{AB}{AB}, \frac{AB}{Ab}, \frac{Ab}{Ab}, \frac{AB}{aB}, \frac{AB}{ab}, \frac{Ab}{aB}, \frac{Ab}{ab}, \frac{aB}{aB}, \frac{aB}{ab}, \frac{ab}{ab}.$$

The physical manifestation of the genotype is called the **phenotype**. If the genotype Aa has the phenotype of the AA individual, then A is said to be a **dominant** gene and a is called **recessive** to A.

We shall assume that an offspring is formed by the donation of a **gamete** (one of each pair of homologous chromosomes) from each of two parents. In the case of one locus, each parent, depending on its genotype, may donate either A or a to form a **zygote** (fertilized egg) having **genotype** AA, Aa or aa. Individuals with genotype AA or aa are **homozygotes**; Aa is a **heterozygote**. For two loci, the donated gametes can be of four kinds, AB, Ab, aB or ab and ten zygotes are possible as listed previously. Generations are taken to be non-overlapping.

Considering the one locus case, we are primarily interested in tracing the frequencies of the three genotypes over time. Assume that the population size is very large,

effectively infinite. Let u_n, v_n, and w_n be the frequencies of AA, Aa and aa, respectively, in the nth generation. In order to follow the vector (u_n, v_n, w_n) as n increases we must describe the mating system, i.e., the way mating pairs are to be selected.

One of the most widely studied systems of mating is **random-mating.** This occurs when any one individual of one sex is equally likely to mate with any one of the opposite sex. Thus, in the one locus case above, the mating $AA \times AA$ would occur with frequency u_n^2 at the nth generation. From this mating only AA offspring result. However, from the mating $Aa \times Aa$, AA, Aa and aa offspring will be produced with probabilities $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ respectively. This equally likely case of segregation is called **Mendelian segregation.**

In an infinite population, not subject to any outside influences, and in which random mating takes place the **Hardy-Weinberg Law** holds. This states that, if in a given generation the frequencies of the A and a gene are p and $q = 1 - p$ respectively, then in all subsequent generations the frequencies remain the same. Verification of this, and the fact that **random mating** is equivalent to **random union of gametes** can be found in most textbooks in population genetics, e.g., Kempthorne [**24**] Chapter 2.

There are a number of factors (apart from the mating system) which act on populations to influence the path of evolution. Perhaps the three most familiar are **mutation, migration** and **selection.** The first two are self-explanatory. They can be visualized as providing the raw material for selection to mould. We are interested here in three forms of selection. The first is selection through variation in **viability,** i.e. the genotypes differ in their chances of survival to reproduce. The second is through **fertility** variations, i.e., different pairs of parents, on account of the genotype of both parents may produce differing numbers of offspring. **Segregation distortion** from the usual Mendelian ratios is another type of selection. These can be considered particular manifestations of what was called by Darwin (1859) "fitness" in his qualitative description of the different abilities of individuals to survive and contribute to the next generation. Of course, the mating system itself can be another factor affecting evolution. Selection attributable to the mating system is commonly referred to as **sexual selection** to distinguish it from **natural selection.** We shall be partly interested in the mathematical description of the interactions between selection and various mating systems.

Selection is incorporated mathematically in the following ways: If the mating type $AA \times Aa$ is assumed to have fertility f then the offspring are produced in the proportions $\frac{1}{2}f\ AA$, $\frac{1}{2}f\ Aa$. Similar definitions hold for the other matings. The offspring are assumed to have viabilities in the ratio $\sigma_1 : \sigma_2 : \sigma_3$ means that each of the genotypes AA, Aa and aa survives to parenthood with relative chance $\sigma_1 : \sigma_2 : \sigma_3$ respectively.

The frequencies u_n, v_n, w_n of AA, Aa, aa in the nth generation can now be expressed in terms of those in the $(n - 1)$-th generation using some transformation T which will in general be non-linear.

Another phenomenon of considerable importance to the maintenance of genetic variability will be mentioned before we describe the models in detail. Recombination may occur in the case of two loci when at the first locus we have alleles A and a and at the second B and b, and the two loci are not independent so far as gamete donation is concerned. An individual heterozygous at both loci can produce four types of gametes. For example, an individual of genotype AB/ab can produce gametes of type AB and ab and also gametes of the type Ab, aB. When all four are produced in equal numbers the loci are called **unlinked.** The AB and ab gametes are called **parental** while the Ab and aB are called **recombinant.** If the loci are linked there will be an excess of parental gametes over recombinants. It is found that the parental types AB, ab are produced with equal frequencies $\frac{1}{2}(1-r)$ and the recombinant types with equal frequencies $\frac{1}{2}r$ where the number $r, 0 < r \leqq 1$, is called the **recombination fraction.** For the physical explanation of the phenomenon and more details on its importance the reader should consult any genetics text book.

This has been a necessarily brief introduction to the terminology we shall use. No attempt has been made to elaborate the biological scope of the terms introduced. For this the reader should consult such texts as Stern [**32**], Crow and Kimura [**7**], and Cavalli and Bodmer [**6**].

II. SOME ONE LOCUS SELECTION MODELS

1. One sex viability model. Consider a population with two possible alleles A, a at a specified locus undergoing random mating and subject to viability selection where the genotypes AA, Aa and aa which survive to maturity (i.e., to reproduce) are in the ratio $\sigma_1 : \sigma_2 : \sigma_3$ respectively.

If the frequencies of A and a in the current generation are p and $q = 1 - p$ respectively, then random union of genes (which is equivalent to random mating) produces the genotypes AA, Aa, aa in the frequencies p^2, $2pq, q^2$ respectively. The relative frequencies of the three genotypes at maturity taking account of selection effects are then

$$\begin{array}{ccc} AA & Aa & aa \\ \sigma_1 p^2 & \sigma_2 2pq & \sigma_3 q^2. \end{array}$$

With Mendelian segregation (see Section I) the frequency p' and q' of A and a respectively, in the next generation have relative magnitudes $p' \sim p^2\sigma_1 + \sigma_2 pq$, $q' \sim \sigma_3 q^2 + \sigma_2 pq$. To convert these to *bona fide* frequencies we normalize by dividing by the sum yielding the transformation equation

$$p' = \frac{p^2\sigma_1 + \sigma_2 pq}{p^2\sigma_1 + 2pq\sigma_2 + q^2\sigma_3} \stackrel{\text{def}}{=} f(p). \tag{2.1}$$

The denominator is commonly called the **mean fitness function**, written $W(p)$, and enjoys the remarkable property that $W(f(p)) \geqq W(p)$ with equality holding iff $p = f(p)$.

The evolution of the process is obtained by iterating the transformation law (2.1). The following classical results are readily established (cf. Figure 1 below) independent of the initial p $(0 < p < 1)$.

$$\lim_{n\to\infty} f_{(n)}(p) = \lim_{n\to\infty} f(f_{(n-1)}(p)) = 1\ (=0) \text{ when } \sigma_1 \geqq \sigma_2 > \sigma_3\ (\sigma_3 \geqq \sigma_2 > \sigma_1), \tag{2.2}$$

$$\lim_{n\to\infty} p_n = \hat{p} = \frac{\sigma_2 - \sigma_3}{2\sigma_2 - \sigma_1 - \sigma_3} \text{ when } \sigma_2 > \max(\sigma_1, \sigma_3). \tag{2.3}$$

In the case $\min(\sigma_1, \sigma_3) > \sigma_2$ then

$$\lim_{n\to\infty} p_n = 1 \text{ for } p > \hat{p}, = 0 \text{ for } p < \hat{p}. \tag{2.4}$$

Figure 1 shows what happens to $f_{(n)}(p)$ in graphical form. The rigorous details are easily supplied.

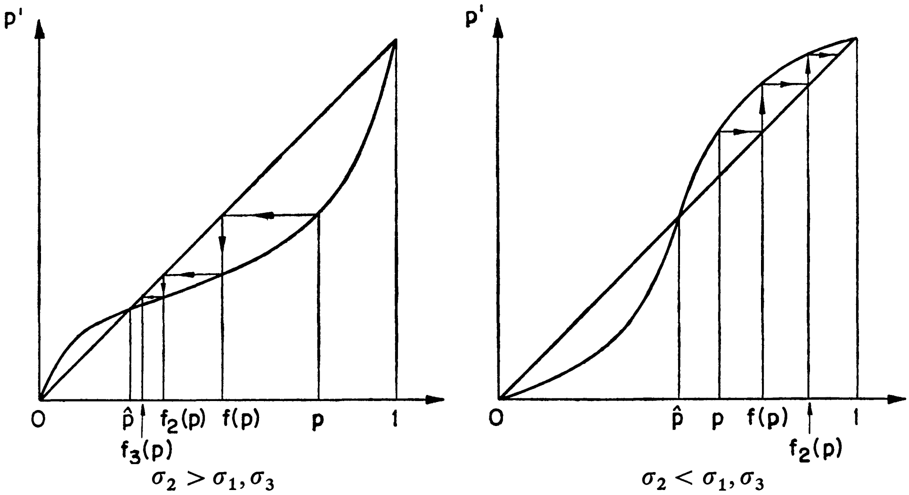


Fig. 1.

The equilibrium $\hat{p}$ is of great importance biologically because it entails the simultaneous existence at an equilibrium involving all genotypes. Thus when the heterozygote is the most fit of the three genotypes a stable **polymorphism** (with all forms) will be maintained. The model of heterozygote advantage (also called the principle of overdominance) has been central to the development of theories on the existence of genetic variability.

2. Two sex viability models with two alleles. (This model was most recently dealt with by Bodmer [2], see also Karlin [20].) Consider next a population divided into males and females, mating randomly subject to viability selection where the fitness coefficients may differ between the sexes. The array in Table 1 describes the process (assuming male and female offspring are produced with equal probability).

Sex		Male			Female	
Gamete		A	a		A	a
Frequency		p	q		P	Q
Genotype	AA	Aa	aa	AA	Aa	aa
Fitness coefficients (viabilities)	σ	1	τ	s	1	t
Relative frequencies after random mating and selection	σpP	$pQ + qP$	τqQ	spP	$pQ + qP$	tqQ

TABLE 1

With Mendelian segregation we obtain for the gene frequencies in the next generation the transformation equations

$$(2.5) \qquad p' = \frac{\sigma pP + \frac{1}{2}(pQ + qP)}{\sigma pP + pQ + qP + \tau qQ}, \qquad P' = \frac{spP + \frac{1}{2}(pQ + qP)}{spP + pQ + qP + tqQ},$$

where the denominators are the required normalization factors (cf. Model 1).

In the case at hand it is more convenient to express the changes of gene frequencies over successive generations in terms of the equivalent pair of variables $x = p/q$, $y = P/Q$, $0 \leqq x, y \leqq \infty$. We obtain

$$(2.6) \qquad x' = \frac{\sigma xy + \frac{1}{2}(x + y)}{\tau + \frac{1}{2}(x + y)} = f(x, y), \qquad y' = \frac{sxy + \frac{1}{2}(x + y)}{t + \frac{1}{2}(x + y)} = g(x, y).$$

Write T for the mapping defined in (2.6). The fixed point $\mathbf{0} = (0,0)$ corresponds to the pure population of only aa genotypes and $\infty = (\infty, \infty)$ represents the pure population of AA genotypes.

We wish to ascertain the character of all equilibria of T and their domains of attraction. The analysis of T and its iterates is much facilitated by exploiting the feature that T is monotone, i.e., where $z = (x, y) \leqq \tilde{z} = (\tilde{z}, \tilde{y})$ holds (the ordering signifies the inequality for each coordinate). Then we have

(2.6a) $\quad Tz \leqq T\tilde{z}$ with strict inequality in each coordinate unless $z = \tilde{z}$.

The stability nature of any equilibrium is customarily ascertained by analysis of the local linear approximation to the non-linear mapping T in the neighborhood of the fixed point. More specifically, we examine the matrix transformation given by the gradient matrix

$$\| \partial T \| = \left\| \begin{matrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial g}{\partial x} \\ \dfrac{\partial f}{\partial y} & \dfrac{\partial g}{\partial y} \end{matrix} \right\|$$

evaluated at the fixed point $\hat{z} = (\hat{x}, \hat{y})$.

If both eigenvalues of $\partial T|_{\hat{z}}$ are in magnitude less than 1, then $\hat{z}$ is locally stable. If at least one eigenvalue in magnitude exceeds 1, then *usually* $\hat{z}$ is unstable.

The conditions for local stability of the pure equilibrium **0** and ∞ are readily determined by invoking the local linear analysis just described. We get

$$\begin{aligned} &\mathbf{0} \text{ (fixation in the } a \text{ gene) is stable iff } \frac{1}{2\tau} + \frac{1}{2t} \leqq 1 \\ (2.7)\qquad & \\ &\infty \text{ (pure } AA \text{ population) is stable iff } \frac{1}{2\sigma} + \frac{1}{2s} \leqq 1. \end{aligned}$$

Algebraic manipulations of the equations (2.6) show that for general positive fitness parameters (σ, τ, s, t) there exist at most 3 fixed points where both coordinates are positive and finite. These are, of course, polymorphic equilibria.

There are five qualitative cases of interest:

(i) *The same homozygote is most fit in both sexes; e.g.,* $\sigma < 1 < \tau$ *and* $s < 1 < t$ *hold.* Under these conditions adding the relations in (2.6) using obvious inequalities produces

$$x' + y' < 2\,\frac{xy + \frac{1}{2}(x+y)}{1 + \frac{1}{2}(x+y)}. \tag{2.8}$$

Since $4xy \leqq (x+y)^2$ we see that $x' + y' < x + y$. It follows that $x^{(n)} + y^{(n)}$ decreases in n and its limit is necessarily zero indicating that **0** is globally stable.

(ii) *AA is most fit in one sex and aa is most fit in the other sex.* We illustrate with the special symmetric situation $\tau = s$ and $\sigma = t$, $\sigma > 1 > s$. In this case there always exists a unique internal equilibrium $z^* = (\xi_0, 1/\xi_0)$ where ξ_0 is the unique positive solution of the equation

$$\xi^3 + \xi^2(2s-1) - \xi(2\sigma - 1) - 1 = 0.$$

Analysis reveals that z^* is stable iff the equilibrium point **0** (and simultaneously, owing to symmetry, the point ∞) is unstable, i.e., iff $1/2\sigma + 1/2s > 1$.

In the general case of (ii) it can be proved that there can be at most one polymorphic stable equilibrium.

(iii) *Both homozygotes selectively inferior to the heterozygote in one sex but superior in the other sex*, i.e.,

$$1 > \sigma, \tau, \qquad 1 < s, t. \tag{2.9}$$

We illustrate with the symmetric case $\sigma = \tau$ and $s = t$. Then $z^* = \mathbf{1} = (1,1)$ is a fixed point of the mapping T and is locally stable iff $\sigma s < 1$. If we determine the values of $\sigma = \tau$, $s = t$ satisfying

$$\frac{2}{\dfrac{1}{s} + \dfrac{1}{\sigma}} < 1 < \sqrt{\sigma s}$$

which is certainly possible (owing to the harmonic mean, geometric mean inequality) we find that **0**, **1** and ∞ are all unstable. Exploiting the monotonic nature of T, we deduce the existence of two other stable polymorphic equilibrium. Here, then, is a case of the existence of two stable polymorphisms. This phenomenon does not arise in the corresponding one sex model.

(iv) *Heterozygote advantage in each sex* $(1 > \sigma, \tau, s, t)$. The expected intuitive result of a unique stable polymorphism is indeed realized.

(v) *Heterozygote advantage in one sex and directed selection in the other sex, i.e.*, $1 > \sigma$, τ, $s > 1 > t$. In this case, elementary analysis of the transformation (2.6) yields the existence of at most two stable equilibria and when two exist one has to be a boundary equilibrium.

To sum up, the main conclusions are as follows:

There can exist at most two stable equilibria including the possibility that both are polymorphisms. In contrast, the one sex selection model allows at most one stable polymorphism.

3. Two sex multi allele viability model. Suppose there exist $r \geqq 3$ alleles $A_1, A_2, \cdots, A_r$ possible at the given locus and of course, $r(r+1)/2$ possible genotypes A_iA_j. Let the frequencies of the genes in the male population be $q_1, q_2, \cdots, q_r$ and $p_1, p_2, \cdots, p_r$ for the female population. The viability fitness matrix for females is designated as $F = \|f_{ij}\|_{i,j=1}^{r}$ where f_{ij} measures the relative average number of the A_iA_j genotype that survive to maturity. The viability fitness matrix for males is denoted by $M = \|m_{ij}\|$.

Stipulating random union of genes and Mendelian segregation quite analogous to (2.5), we obtain for the gene frequencies of the next generation the recursion relations

$$(2.10)\qquad \begin{aligned} p_i' &= \frac{\frac{1}{2}\left[p_i \sum_{j=1}^{r} f_{ij}q_j + q_i \sum_{j=1}^{r} f_{ij}p_j\right]}{\sum_{i,j=1}^{r} p_i f_{ij} q_j} \\ q_i' &= \frac{\frac{1}{2}\left[p_i \sum_{j=1}^{r} m_{ij}q_j + q_i \sum_{j=1}^{r} m_{ij}p_j\right]}{\sum_{i,j=1}^{r} p_i m_{ij} q_j}, \qquad i = 1, 2, \cdots, r. \end{aligned}$$

Call this non-linear transformation of $2r$ variables ($2r - 2$ independent ones) T as before. Results concerning the evolution of this process, i.e., the behavior of the iterates of T and characterizing their limit points, are of primary interest. It would be of much interest to determine precise bounds for the number of stable polymorphisms possible in this r allele selection model. Theorems from algebraic geometry produce upper bounds (but excessive ones) for the number of admissible equilibrium points. We refer to Karlin [**20**] for a treatment of several non-elementary

cases of (2.10). A rather complete treatment of the special symmetric case $M = F$ is available, e.g., see Kingman [**25**].

4. Selection model for multi allelic sex linked character. (This model was first formulated by Haldane, see also Cannings [**4**], [**5**].)

Consider a character determined by a locus on the sex chromosome with r alleles possible. Suppose the female sex is the homogametic one, the XX chromosome.

The female genotypes assume the form A_iA_j, $i,j = 1,\cdots,r$ but the male genotypes take the form A_iY since the Y chromosome carries no complement of the gene.

The fitness coefficients corresponding to females are displayed by the matrix $F = \|f_{ij}\|$ and for males by the vector $\boldsymbol{m} = (m_1, m_2, \cdots, m_r)$. Thus m_i measures the relative fitness of the male genotype A_iY and f_{jk} of the female genotype A_jA_k. Under random mating and selection, the relative number of female offspring of type A_jA_k which survive to maturity is $\frac{1}{2}(p_jq_k + q_jp_k)f_{jk}$ for $j \neq k$ and $p_jq_jf_{jj}$ for $j = k$. For males of genotype A_i, the relative frequency of maturing male offspring is q_im_i, since the male parent always contributes the Y chromosome. With Mendelian segregation, we get the transformation law

$$p_i' = \frac{\frac{1}{2}\left[p_i \sum_{j=1}^{r} f_{ij}q_j + q_i \sum_{j=1}^{r} f_{ij}p_j\right]}{\sum_{i,j=1}^{r} p_if_{ij}q_j}, \quad q_i' = \frac{m_ip_i}{\sum_{i=1}^{r} m_ip_i}. \tag{2.11}$$

In general, there exists at most one polymorphic equilibrium $\hat{p}$, $\hat{q}$ where $\hat{p}$ is calculated by normalizing (so that the sum of components is 1) the positive solution of

$$(FI_m + I_mF)\hat{p} = \mathbf{1}. \tag{2.12}$$

(I_m is the diagonal matrix with $m_1, m_2, \cdots, m_r$ down the diagonal and $\mathbf{1}$ is the vector with all components of value 1.) And

$$\hat{q} = \gamma I_m\hat{p} \quad \text{with } \gamma^{-1} = \sum_{i=1}^{r} m_i\hat{p}_i.$$

Stability conditions of such a polymorphic solution can be determined.

We specialize now to the case $r = 2$. Then it is more convenient to work in terms of the variables

$$x = \frac{p_1}{p_2} \quad \text{and } y = \frac{q_1}{q_2},$$

so that $0 \leqq x, y \leqq \infty$. The equivalent recursion equations reduce to

$$x' = \frac{sxy + \frac{1}{2}(x + y)}{\sigma + \frac{1}{2}(x + y)}, \qquad y' = mx, \tag{2.13}$$

where $s = f_{11}/f_{12}$, $\sigma = f_{22}/f_{12}$ and $m = m_1/m_2$. Designate the transformation (2.13) as $T(x, y) = (x', y')$. It is readily verified that T is a strictly monotonic mapping

(cf. (2.6a)). Exploiting this fact we easily establish by applying a local linear approximation, the existence of a positive pair of numbers (a, b) such that for $\varepsilon > 0$ and sufficiently small $T(\varepsilon a, \varepsilon b) < (\varepsilon a, \varepsilon b)$ iff $m < 2\sigma - 1$. It follows that the fixed point $\mathbf{0} = (0,0)$ (corresponding to a pure A_2A_2 population) is locally stable iff $m \leqq 2\sigma - 1$. In a similar manner, we find that $\infty = (\infty, \infty)$ is locally stable iff $2s - 1 \geqq 1/m$. For the case where $2\sigma - 1 < m$ and $2s - 1 < 1/m$ there exists a unique polymorphic globally stable equilibrium (x^*, y^*) with

$$x^* = \frac{2\sigma - 1 - m}{(2s-1)m - 1}, \qquad y^* = mx^*. \tag{2.14}$$

Global stability of (x^*, y^*) results by virtue of the following facts: (i) T is monotone and exactly one interior equilibrium exists, (ii) $T(\varepsilon a, \varepsilon b) > (\varepsilon a, \varepsilon b)$, and (iii) $T(N\tilde{a}, N\tilde{b}) < (N\tilde{a}, N\tilde{b})$ hold for ε small enough and N large enough respectively. (Here a, b are specified to satisfy $m > a/b > 2\sigma - 1$ and $\tilde{a}, \tilde{b}$ to satisfy $1/m > \tilde{a}/\tilde{b} > 2s - 1$.)

In the case that $m < 2\sigma - 1$ and $2s - 1 < 1/m$ simultaneously hold then $\mathbf{0}$ and ∞ are both locally stable and possess domains of attraction whose boundary is an algebraic curve containing the point (x^*, y^*) defined in (2.14).

5. Segregation distortion and viability selection balance for the *t*-locus in house mice. (This model was set up by Lewontin [**28**].)

The t-locus codes for certain enzyme function essentially involves two alleles labeled T and t. The presence of the t-alleles affects males and females differently. (Morphologically the t allele reveals a shortened tail—hence the name.) With reference to selection, we have

	Male			Female		
	TT	Tt	tt	TT	Tt	tt
Fitnesses	$1 - s$,	1,	0	$1 - s$	1	$1 - \sigma$

$(0 \leqq s < 1, 0 \leqq \sigma \leqq 1)$. Note that recessive males (tt genotypes) suffer total lethality.

The main difference is revealed in the segregation ratios for the heterozygote in the two sexes. Explicitly

	Males		Females	
	$T \swarrow Tt \searrow t$		$T \swarrow Tt \searrow t$	
	T	t	T	t
segregation ratios	$1 - m$	m	$\frac{1}{2}$	$\frac{1}{2}$

and m is about .90 in the actual example.

Denote by q_1 (q_2) the frequency of $T(t)$ in the males and $p_1(p_2)$ correspondingly for females. Set $u = q_2/q_1$, $v = p_2/p_1$. Taking account of the viability selection, segregation bias and assuming random mating, we deduce the recursion relations

$$(2.15)\qquad u' = \frac{(1-\sigma)uv + \frac{1}{2}(u+v)}{1-s+\frac{1}{2}(u+v)}, \qquad v' = \frac{m(u+v)}{1-s+(1-m)(u+v)}.$$

The transformation (2.15) is strictly *monotonic* as in the earlier two allele models. Direct examination reveals that the transformation Γ in (2.15) satisfies

$$(2.16)\qquad \Gamma(\varepsilon a, \varepsilon b) > (\varepsilon a, \varepsilon b)$$

for $\varepsilon > 0$ small enough and appropriate $a, b > 0$ iff $2(1-s)\ (\frac{1}{2} - m - s) < 0$ or $m + s > \frac{1}{2}$ and the opposite order relation holds in (2.16) when $m + s < \frac{1}{2}$.

It follows that $\mathbf{0} = (0,0)$ is locally stable iff $m + s \leqq \frac{1}{2}$. We now prove global stability for this case. To this end form

$$u' + v' \leqq \frac{(1-\sigma)uv + \frac{1}{2}(u+v)}{1-s+\frac{1}{2}(u+v)} + \frac{m(u+v)}{1-s+(1-m)(u+v)}$$

$$\leqq \frac{(1-\sigma)uv + (m+\frac{1}{2})(u+v)}{1-s+\frac{1}{2}(u+v)}.$$

But $uv \leqq ((u+v)/2)^2$ implies

$$(2.17)\qquad z' = u' + v' \leqq \frac{(m+\frac{1}{2})z + (1-\sigma)z^2/4}{1-s+\frac{1}{2}z} = h(z).$$

Direct verification shows that h is non-decreasing and $h(z) \leqq z$ for $z \geqq 0$ with equality iff $z = 0$. Iteration of (2.17) is therefore permissible leading to

$$z^{(n)} \leqq h_n(z) = h(h_{n-1}(z)), \qquad n = 1, 2, 3, \cdots$$

But a simple geometric argument proves $h_n(z) \to 0$ as $n \to \infty$ for any initial $z > 0$ and therefore $z^{(n)} \to 0$. Thus $\mathbf{0} = (0,0)$ is globally stable as claimed.

The fixed points of (2.15) are obtained as the solutions of the equations

$$(2.18)\qquad u = \frac{(1-s-m)v + (1-m)v^2}{m - (1-m)v},$$

where v satisfies $R(v) = A_3v^3 + A_2v^2 + A_1v + A_0 = 0$, where

$$(2.19)\qquad \begin{aligned} A_0 &= m(1-s)(s+m-\tfrac{1}{2}), \\ A_1 &= m(1-\sigma)(1-s-m) + (1-s)[-2m(1-m) + s(m-\tfrac{1}{2})], \\ A_2 &= (1-m)[(1-\sigma)(2m+s-1) - (1-s)(m-\tfrac{1}{2})], \\ A_3 &= -(1-\sigma)(1-m)^2. \end{aligned}$$

When $s + m > \frac{1}{2}$ we have $R(0) > 0$ while

$$R\left(\frac{m}{1-m}\right) = \frac{-m}{2}\ \frac{(1-s)^2}{1-m} < 0.$$

Therefore, in this case there exists v^* ($0 < v^* < m/(1-m)$) satisfying $R(v^*) = 0$ and u^* determined from (2.18) is > 0. The point (u^*, v^*) is of course an equilibrium of (2.15). With a little effort, using σ as a parameter ($1 \geqq \sigma \geqq 0$) it can be proved there exists for $m + s > \frac{1}{2}$ a unique solution v^* of $R(v) = 0$ fulfilling the inequalities $0 < v^* < m/(1-m)$ and therefore in this case exactly one interior polymorphism occurs. Since $T(Na, Nb) < (Na, Nb)$ prevails for N large enough and appropriate $a > 0$, $b > 0$, we infer, by virtue of the monotonic nature of T the limit relation $\lim_{n\to\infty} T^n(u,v) = (u^*, v^*)$ from any initial $(u,v) > 0$.

6. Another model of segregation distortion. We close this section by citing a one-locus two allele segregation distortion model considered by Haldane [**14**]. There are no fertility differences in the mating types or viability selection differences. There are two alleles A_1 and A_2 where the frequencies of A_1A_1, A_1A_2 and A_2A_2 are x, y and z respectively. The array in Table 2 decribes the segregation ratios depending on two parameters.

Mating	Offspring ratios A_1A_1	A_1A_2	A_2A_2	Mating Frequency
$A_1A_1 \times A_1A_1$	1	0	0	x^2
$A_1A_1 \times A_1A_2$	λ	$1-\lambda$	0	$2xy$
$A_1A_1 \times A_2A_2$	0	1	0	$2xz$
$A_1A_2 \times A_1A_2$	$\dfrac{\lambda(1-\mu)}{2-\lambda-\mu}$	$\dfrac{2(1-\lambda)(1-\mu)}{2-\lambda-\mu}$	$\dfrac{\mu(1-\lambda)}{2-\lambda-\mu}$	y^2
$A_1A_2 \times A_2A_2$	0	$1-\mu$	μ	$2yz$
$A_2A_2 \times A_2A_2$	0	0	1	z^2

TABLE 2.

Viability effects only operate in the segregation process. Each mating has output 1. It is straightforward to derive the recursion relations connecting genotype frequencies over two successive generations. We get

$$\begin{aligned} x' &= x^2 + 2\lambda xy + \frac{\lambda(1-\mu)}{2-\lambda-\mu}\, y^2 \\ y' &= 2(1-\lambda)xy + 2xz + \frac{2(1-\lambda)(1-\mu)}{2-\lambda-\mu} y^2 + 2(1-\mu)yz \qquad (2.20) \\ z' &= z^2 + 2\mu yz + \frac{\mu(1-\lambda)}{2-\lambda-\mu} y^2. \end{aligned}$$

All equilibria can be determined in general, and for some special cases, viz., $\lambda = \mu$, $\lambda = 1 - \mu$, $\lambda = 0$ or 1, the full convergence behavior can be analysed.

Thus, when $\mu = 0$, $x^{(n)} \to 1$ rapidly.

When $\lambda + \mu = 1$ and $\lambda > \frac{1}{2}$, again we find $x^{(n)} \to 1$.

For $\lambda = \mu$ and $\lambda < \frac{1}{2}$, then it can be proved that

$$x^{(n)}; z^n \to \frac{1 - \sqrt{1 - 2\lambda(1 - 2\lambda)}}{2(1 - 2\lambda)}.$$

The following can be readily checked. Assume by symmetry $(0 < \mu \leqq \lambda < 1)$ then:

(i) For $0 < \mu \leqq \lambda < \frac{1}{2}$, there exists a unique locally stable polymorphism.

(ii) For $0 < \mu < \frac{1}{2} < \lambda < 1$, there exists no internal equilibrium. It can be proved that fixation in the A_1 allele occurs.

(iii) If $\frac{1}{2} < \mu \leqq \lambda < 1$, there exists a unique internal non-stable equilibrium.

The global convergence behavior of (2.20) for arbitrary parameters λ, μ is in general unsettled.

III. SOME MODELS OF POSITIVE ASSORTATIVE MATING

Consider a two-allele (A and a) single locus population displaying certain preferences in mating behavior. We consider here the case where the preference is exercised by one of the sexes, say the female sex, (this covers most situations of insect and mammal populations). (References and more detailed discussion of the models and related models of this section can be found in Scudo and Karlin **[30]** and Karlin and Scudo **[18]**.)

1. A model of assortative mating. Assume that A is dominant to a so that phenotypically AA and Aa are alike. The degree of partial assortative mating in the phenotypes is measured by two parameters: α $(0 \leqq \alpha \leqq 1)$ will be the fraction of dominant females preferring to mate with their own kind and β $(0 \leqq \beta \leqq 1)$ that of recessive females preferring their own kind. Thus a fraction, $1 - \alpha$, of $\bar{A}$ (of AA or Aa) females mate indifferently, i.e., at random. We assume all females are fertilized (i.e., find a suitable mate). This happens if the males are sufficiently abundant and the same male may participate in many matings. Consider the genotypes AA, Aa, aa (A dominant) with the frequencies u, v and w respectively in the female population.

When the prohibitions of assortative mating are operating, it is obligate that each mate of an aa individual is of the same genotype so that the frequency of the $aa \times aa$ mating type is w. Therefore the frequency of the matings of the dominant phenotypes is $1 - w = u + v$. Among the matings of dominants the frequency of occurrence of the $AA \times AA$ mating type is u^2 and its frequency of occurrence considering all admissible matings is then $u^2/(1 - w)$. The frequencies of the mating types are listed in Table 3.

Mating Type	Frequencies	
	Of Assorting Types	Random Mating
$AA \times AA$	$\alpha u^2/(u+v)$	$(1-\alpha)u^2$
$AA \times Aa$	$2\alpha uv/(u+v)$	$2(1-\alpha)uv$
$AA \times aa$		$(2-\alpha-\beta)uw$
$Aa \times Aa$	$\alpha v^2/(u+v)$	$(1-\alpha)v^2$
$Aa \times aa$		$(2-\alpha-\beta)vw$
$aa \times aa$	βw	$(1-\beta)w^2$

TABLE 3

The corresponding recurrence relations connecting genotype frequencies over successive generations in accordance with Mendelian segregation laws become

$$u' = \left(\frac{\alpha}{u+v} + (1-\alpha)\right)(u+\tfrac{1}{2}v)^2$$

$$(3.1) \qquad v' = \frac{\alpha uv}{2(u+v)} + \tfrac{1}{2}v + (1-\alpha)u(\tfrac{1}{2}v+w) + (1-\beta)w(u+\tfrac{1}{2}v),$$

$$w' = \beta w + \frac{\alpha v^2}{4(u+v)} + (1-\alpha)\tfrac{1}{2}v(\tfrac{1}{2}v+w) + (1-\beta)w(\tfrac{1}{2}v+w).$$

Introducing the A gene frequency, $p = u + \frac{1}{2}v$, and for the next generation, $p' = u' + \frac{1}{2}v'$ and, letting p_n denote the frequency of the gene A in the nth generation, we derive, from (3.1), the relationship

$$(3.2) \qquad p' = p[1 + \tfrac{1}{2}(\alpha-\beta)w].$$

The following inferences can now be made:

(i) For $\alpha > \beta$, p_n increases to 1, the pure homozygous AA state. The rate of convergence is algebraic.

(ii) For $\alpha < \beta$, the population ultimately fixes in the pure homozygous aa state and convergence occurs with an asymptotic factor of decrease per generation $\lambda = 1 + \frac{1}{2}(\alpha-\beta)$.

When $\alpha = \beta$ it is readily checked that $p^{(n)} = p^{(0)}$ for all n. Then v' simplifies to

$$v' = \frac{vp\alpha}{p+\frac{1}{2}v} + (1-\alpha)\,2pq = f(v), \qquad (q = (1-p)),$$

where p is the constant gene frequency. Thus $f(v)$ is a linear fractional transformation and therefore the nth generation frequencies $v_n = f_n(v_0) = f(f_{n-1}(v_0))$ can be explicitly evaluated. Indeed, we have

$$\frac{v_n-\gamma_1}{v_n-\gamma_2}=K^n\left(\frac{v_0-\gamma_1}{v_0-\gamma_2}\right),$$

where γ_1 and γ_2 are the fixed points of $f(v)=v$ and

$$K=\frac{\gamma_2}{\gamma_1}\left[\frac{2(1-\alpha)pq-\gamma_1}{2(1-\alpha)pq-\gamma_2}\right].$$

Because $f(v)$ is concave increasing, we deduce $v_n \to \gamma_1$. For the case $\alpha = 1$ we obtain $v_n = 2pv_0/(nv_0+2p)$ so that $v_n \to 0$ at an algebraic rate.

2. Model of assortative mating with permanent bonding. In the formulation of the previous model it was tacitly assumed that there was no set order in which the types of mating (random or assortative) took place. The factor of timing of mating for assorting and random mating individuals may be important, and could affect the accessibility and availability of proper mates.

Two simple contrasting assumptions can be made to study the effect of assortment on the timing of pair bonding depending on whether assorting females mate prior to the nonassorting ones, or after. Let u, v, w denote the frequencies of the AA, Aa and aa genotypes respectively.

In the first set up a fraction $\alpha(u+v)$ of the dominant females pair first with an equal number of dominant males; the same occurs for βw of the recessives. The remaining individuals, a proportion $(1-\alpha)(u+v)+(1-\beta)w$ of both sexes mate at random. The resulting relative frequencies of the mating types are given in Table 4.

Mating Types	Frequencies	
	Assorting	Random Mating
$AA \times AA$	$\alpha \dfrac{u^2}{u+v}$	$(1-\alpha)^2u^2/R$
$AA \times Aa$	$2\alpha \dfrac{uv}{u+v}$	$2(1-\alpha)^2uv/R$
$AA \times aa$		$2(1-\alpha)(1-\beta)uw/R$
$Aa \times Aa$	$\alpha \dfrac{v^2}{u+v}$	$(1-\alpha)^2v^2/R$
$AA \times aa$		$2(1-\alpha)(1-\beta)vw/R$
$aa \times aa$	βw	$(1-\beta)^2w^2/R$

TABLE 4.

One can "normalize" back to frequencies (Case A) simply by dividing the proportions in the random mating part by $(1-\alpha)(u+v)+(1-\beta)w$. On the other hand, we can assume (Case B) that the delay in pairing causes some decrease in reproduction. One way to express the loss in fertility is to assume that the contribution to the next generation on the part of the population undergoing random mating is

$$[(1-\alpha)(u+v)+(1-\beta)w]^2 \text{ instead of } (1-\alpha)(u+v)+(1-\beta)w.$$

An alternative formulation in which random mating females pair first can be analyzed (see Scudo and Karlin [30]).

Recurrence relations for genotype frequencies over successive generations are as follows:

CASE A. $R = 1-\alpha+(\alpha-\beta)w$

$$\begin{aligned} u' &= \alpha\frac{(u+\frac{1}{2}v)^2}{u+v}+(1-\alpha)^2(u+\tfrac{1}{2}v)^2/R, \\ v' &= \alpha v\frac{u+\frac{1}{2}v}{u+v}+2(1-\alpha)(u+\tfrac{1}{2}v)\left\{\frac{1-\alpha}{2}v+(1-\beta)w\right\}\Big/R, \\ w' &= \beta w+\alpha\frac{v^2}{4(u+v)}+\left\{\frac{1-\alpha}{2}v+(1-\beta)w\right\}^2\Big/R. \end{aligned} \tag{3.3}$$

CASE B.

$$\begin{aligned} Nu' &= \alpha\frac{(u+\frac{1}{2}v)^2}{u+v}+(1-\alpha)^2(u+\tfrac{1}{2}v)^2, \\ Nv' &= \alpha v\frac{u+\frac{1}{2}v}{u+v}+2(1-\alpha)(u+\tfrac{1}{2}v)\left\{\frac{1-\alpha}{2}v+(1-\beta)w\right\}, \\ Nw' &= \beta w+\alpha\frac{v^2}{4(u+v)}+\left\{\frac{1-\alpha}{2}v+(1-\beta)w\right\}^2, \end{aligned} \tag{3.4}$$

where $N = 1-R(1-R)$.

From (3.3) it follows that the gene frequency $p = u+\frac{1}{2}v$ is invariant over time, i.e., $p' = p$. Using this fact we can rewrite the second equation of (3.3) in the form

$$v' = \frac{\alpha pv}{p+\frac{1}{2}v}+\frac{2p(1-\alpha)[(1-\beta)q+\frac{1}{2}v(\beta-\alpha)]}{1-\alpha p-\beta q+(\beta-\alpha)\frac{1}{2}v} = f(v), \qquad (q = 1-p).$$

The frequency of Aa in the nth generation is therefore $v_n = f_n(v_0)(f_n(v) = f_{n-1}(f(v)))$. By direct verification we find that $f(v)$ is concave and $f(0) > 0$. It follows that $f(v) = v$ admits a unique solution v^* in $(0,1)$ and, independent of the initial frequency v_n, converges to v^*. The equilibrium v^* depends on p and is computed as the unique root in $(0,1)$ of the cubic

$$- v^3(\beta - \alpha) - 2(1 - \alpha p - \beta q)v^2 + 4v(1 - \beta)(1 - \alpha)p^2 + 8p^2q(1 - \alpha)(1 - \beta) = 0.$$

We turn to the analysis of case **B**. Combining appropriately the equations of (3.4) we obtain

$$p' = p\left[\frac{1 - (1 - \alpha)(1 - R)}{1 - R(1 - R)}\right], \tag{3.5}$$

where $R = 1 - \alpha + (\alpha - \beta)w$. Observe that the multiplying factor of p exceeds 1 (is smaller than 1) if and only if $\alpha > \beta$ ($\alpha < \beta$) independent of w ($0 < w < 1$). We deduce easily the following results.

If $\alpha > \beta$, $p_n \uparrow 1$ as $n \to \infty$, i.e., the population fixes in the homozygote AA state. If $\alpha < \beta$, $p_n \downarrow 0$ as $n \to \infty$.

3. Assortative mating with no dominance. A general formulation of a model of assortment and random mating would involve 9 parameters. Let α_1, α_2 and α_3 ($0 \leqq \alpha_i \leqq 1$, $\alpha_1 + \alpha_2 + \alpha_3 \leqq 1$) be measures of the tendency of an AA female to choose an AA, Aa or aa mate respectively. Then $1 - \alpha_1 - \alpha_2 - \alpha_3$ is a measure of ambivalence in the choice of a mate (mates of random). The parameters α_2 and α_3 can be interpreted as propensities of partial disassortment. Similarly, we denote by $\beta_1, \beta_2, \beta_3$ and $1 - \beta_1 - \beta_2 - \beta_3$ the degrees of assortment and random mating respectively for an Aa female. The aa genotype has corresponding assortment parameters γ_1, γ_2 and γ_3. To illustrate, we discuss the case where all parameters of disassortment are zero, i.e., $\alpha_2 = \alpha_3 = 0$, $\beta_1 = \beta_3 = 0$, and $\gamma_1 = \gamma_2 = 0$ (for simplicity we drop the subscript and write $\alpha_1 = \alpha$, $\beta_2 = \beta$, $\gamma_3 = \gamma$).

Let the frequencies of AA, Aa and aa in the present generation be u, v and w respectively. We assume random mating occurs first, followed by assortative mating. Permanent pairing is assumed and this entails that at the culmination of random mating a total frequency of $\alpha u + \beta v + \gamma w$ males are available to mate with assorting females. Thus the fractions of male and female individuals available for isogenotypic pairings are shown in Table 5.

Genotypes	Proportions of	
	Available Males	Assorting Females
AA	$u(\alpha u + \beta v + \gamma w)$	αu
Aa	$v(\alpha u + \beta v + \gamma w)$	βv
aa	$w(\alpha u + \beta v + \gamma w)$	γw

TABLE 5.

Assorting continues until all possible pairs are formed; the remaining individuals do not contribute to the next generation. Observe that all AA assorting females are

fertilized, if and only if $\alpha u \leqq u[\alpha u + \beta v + \gamma w]$ or, what is the same, $\alpha \leqq (\beta v + \gamma w)/(v + w)$. If we make the simplifying assumption $\gamma = \alpha$, then if $\gamma = \alpha < \beta$ holds, we find that all AA and aa assorting females can pair. The fraction of unfertilized Aa females is $(\beta - \alpha)v(1 - v)$. Verification of the entries in the Table 6 should now be clear.

Mating Types	Frequencies	
	Random Mating	Assortative Mating
$AA \times AA$	$(1-\alpha)\,u^2$	αu
$AA \times Aa$	$(1-\alpha)\,uv + (1-\beta)\,uv$	
$AA \times aa$	$(1-\alpha)\,2uw$	
$Aa \times Aa$	$(1-\beta)\,v^2$	$[\alpha(1-v)+\beta v]\,v$
$Aa \times aa$	$(1-\alpha)\,wv + (1-\beta)\,wv$	
$aa \times aa$	$(1-\alpha)\,w^2$	αw

TABLE 6

The associated recursion relations connecting genotype frequencies over two successive generations are

$$
\begin{aligned}
Nu' &= \alpha u + f\frac{v}{4} + (1-\alpha)u\left(u + \frac{1}{2}v\right) + (1-\beta)\,\frac{1}{2}v\left(u + \frac{1}{2}v\right),\\
Nv' &= f\frac{v}{2} + (1-\alpha)\left[\frac{1}{2}v(1-v) + 2uw\right] + (1-\beta)\,\frac{1}{2}v,\\
Nw' &= \alpha w + f\frac{v}{4} + (1-\alpha)w\left(w + \frac{1}{2}v\right) + (1-\beta)\,\frac{1}{2}v\left(w + \frac{1}{2}v\right),
\end{aligned}
\tag{3.6}
$$

where $N = 1 - (\beta - \alpha)v(1 - v)$ and $f = \alpha(1 - v) + \beta v$.

From (3.6) we have

$$u' - w' = (u - w)\left[\frac{1 - (\beta - \alpha)\frac{1}{2}v}{1-(\beta-\alpha)v(1-v)}\right]$$

so that $|u' - w'| > |u - w|$ if and only if $v < \frac{1}{2}$. Moreover, we always have $(u' - w')(u - w) > 0$. Now

$$v' = \frac{\frac{1}{2}v[\alpha + (\beta - \alpha)v] + (1 - \alpha)[\frac{1}{2}v(1 - v) + 2uw] + (1 - \beta)\frac{1}{2}v}{1 - (\beta - \alpha)v(1 - v)}$$

and therefore since $4uw \leqq (1 - v)^2$ we have

$$v' \leqq \frac{\frac{1}{2}v[\alpha + (\beta - \alpha)v] + \frac{1}{2}(1 - \alpha)(1 - v) + (1 - \beta)\frac{1}{2}v}{1 - (\beta - \alpha)v(1 - v)} = g(v),$$

for all $0 \leqq v \leqq 1$. Direct computation affirms that $g'(v) \geqq 0$ $(0 \leqq v \leqq 1)$. It follows that $v_n \leqq g_n(v) = g_{n-1}(g(v))$ where v_n is the frequency of Aa in the nth generation. The theory of iteration of functions tells us that $g_n(v)$ converges as $n \to \infty$ to the unique fixed point v^* of $g(v) = v$ in $(0,1)$. We find that v^* satisfies

$$(\beta - \alpha)v^3 - \frac{3}{2}(\beta - \alpha)v^2 + v\left(1 - \alpha + \frac{\beta}{2}\right) - \frac{1-\alpha}{2} = 0;$$

examination reveals that $v^* < \frac{1}{2}$. Therefore, for n sufficiently large, it follows that $v_n < \frac{1}{2}$ which implies that ultimately $|u_n - w_n|$ continually increases. Its limit is necessarily one. Combining these facts we have established:

(i) If $u_0 > w_0$ then $u_n \to 1$, $v_n \to 0$, $w_n \to 0$.
If $u_0 < w_0$ then $u_n \to 0$, $v_n \to 0$, $w_n \to 1$.

The approach of v_n to 0 is geometrically fast at the rate $1 - \beta/2$.

(ii) When $u_0 = w_0$, then $v_n \to v^*$ and $u_n = w_n \to (1 - v^*)/2$ at the geometric rate $|g'(v^*)|$.

The analysis when $\alpha = \gamma > \beta$ paraphrases that above. The conclusions are the same as before, except that now v^* is the solution in $(0,1)$ of the cubic

$$(\alpha - \beta)v^3 - (\alpha - \beta)v^2 + v(1 - \tfrac{1}{2}\alpha) - \tfrac{1}{2}(1 - \alpha) = 0.$$

4. Assortative mating preceding random mating, permanent bonding. Here, assortment is assumed to occur first with permanent pairing. The remaining genotypic proportions of AA, Aa and aa individuals practicing random mating is $(1-\alpha)u$, $(1-\beta)v$, $(1-\gamma)w$ respectively. Two cases can be considered according to whether males possess infinite fertility or not. Case B implies a loss of frequency of mating types per generation of magnitude $\alpha u + \beta v + \gamma w$ while Case A assumes no impairment of fertility for females mating randomly. The consequences of the matings are summarized in the recursion relations.

CASE A. $R = 1 - \alpha u - \beta v - \gamma w$, $N = 1$.

CASE B. $R = 1$, $N = 1 - R^*(1 - R^*)$, $R^* = 1 - \alpha u - \beta v - \gamma w$.

$$\begin{aligned} Nu' &= \alpha u + \tfrac{1}{4}\beta v + [(1-\alpha)u + \tfrac{1}{2}(1-\beta)v]^2/R, \\ Nv' &= \tfrac{1}{2}\beta v + 2[(1-\alpha)u + \tfrac{1}{2}(1-\beta)v][(1-\gamma)w + \tfrac{1}{2}(1-\beta)v]/R, \\ Nw' &= \gamma w + \tfrac{1}{4}\beta v + [(1-\gamma)w + \tfrac{1}{2}(1-\beta)v]^2/R. \end{aligned} \tag{3.7}$$

We treat only Case B (see Karlin and Scudo [**18**] for case A).

In the present discussion we restrict attention to the important case where $\alpha = \gamma$. We obtain from (3.7)

$$u' - w' = (u - w)\left[\frac{1 - (1-\alpha)(1-R^*)}{1 - R^*(1-R^*)}\right]. \tag{3.8}$$

It follows that $|u' - w'| < |u - w|$ if $\alpha < \beta$ and the opposite inequality holds when $\alpha > \beta$ provided $v > 0$. The recursion relations (3.7) admit a single polymorphic equilibrium $(\hat{u}, \hat{v}, \hat{w})$ where $\hat{w} = \hat{u} = (1 - \hat{v})/2$, and $\hat{v}$ is the unique root in $(0,1)$ of the equation

$$(3.9)\qquad \begin{aligned}(\alpha - \beta)^2 v^3 &+ v^2(\alpha - \beta)[1 - 5/2\alpha + \tfrac{1}{2}\beta] \\ &+ v[1 - \alpha(1 - \alpha) + \tfrac{1}{2}\beta - (\alpha - \beta)(1 - \alpha)] - \tfrac{1}{2}(1 - \alpha)^2 = 0.\end{aligned}$$

(i) When $\alpha > \beta$, it can be easily proved that fixation ultimately occurs.

(ii) When $0 < \alpha < \beta$, then for any nontrivial initial values u_0, v_0, w_0 the genotype frequencies at the nth generation u_n, v_n, w_n converge as $n \to \infty$ to the stable polymorphic equilibrium $(\hat{u}, \hat{v}, \hat{w})$ at a geometric rate. The following is a sketch of the proof.

From (3.8) for the case at hand, we deduce that $u_n - w_n \to 0$. The second relation of (3.1) can be written in the form

$$(3.10)\qquad v_{n+1} = \frac{\tfrac{1}{2}\beta v_n + (1 - \alpha)(1 - \beta)v_n(1 - v_n) + \tfrac{1}{2}(1 - \beta)^2 v_n^2 + (1 - \alpha)^2[\tfrac{1}{4}(1 - v_n)^2 - \tfrac{1}{2}(u_n - w_n)^2]}{1 - R_n^*(1 - R_n^*)}$$

where $R_n^* = 1 - \alpha + (\alpha - \beta)v_n$.

We regard $\tfrac{1}{2}(u_n - w_n)^2 = \varepsilon_n$ as a parameter, and the transformation then achieves the form

$$(3.11)\qquad v_{n+1} = f_\varepsilon(v_n),$$

where $f_\varepsilon(v)$ is the function of (3.10) with ε_n replaced by ε. Simple analysis shows that $f_0(v)$ on $(0,1)$ is monotone increasing and crosses the 45° line at the unique root of the cubic (3.9). Furthermore, $f_\varepsilon(v)$ is monotone increasing for $\eta(\varepsilon) < v < 1 - \eta(\varepsilon)$ with $\eta(\varepsilon)$ tending to zero as $\varepsilon \to 0$.

Inspection of (3.10) reveals that v_n is bounded away from 0 and 1 provided $0 < v_0 < 1$. We infer from (3.11) that

$$f_0^{(m)}(v_{n_0}) \geqq v_{n_0+m} \geqq f^{(m)}(v_{n_0}),$$

where $f_\varepsilon^{(m)}$ denotes the mth composed function f_ε with itself. Letting $m \to \infty$ and exploiting the cited monotonicity properties of f_ε we find that

$$\varliminf_{m\to\infty} v_m \geqq \hat{v}^{(\varepsilon)} \quad\text{and}\quad \varlimsup_{m\to\infty} v_m \leqq \hat{v},$$

where $\hat{v}_\varepsilon$ is the unique fixed point of $f_\varepsilon(v) = v$ in $(\eta(\varepsilon),\ 1 - \eta(\varepsilon))$. Obviously $\hat{v}^{(\varepsilon)} \to \hat{v}$ as $\varepsilon \to 0$ and thus the convergence of v_n to $\hat{v}$ is established. The convergence $u_n \to \tfrac{1}{2}(1 - \hat{v})$ and $w_n \to \tfrac{1}{2}(1 - \hat{v})$ readily ensue.

For the case of general parameters α, β, γ a complete analysis as above appears difficult; however, investigation of local stability of the fixations provides a good

qualitative picture of the properties of the system (3.7). (See Karlin and Scudo [**18**] for details.)

5. Partial assortative mating with no priorities. We now consider the case of mixed assortative and random mating where the two mating patterns occur in no predetermined order. Enough males are assumed to be present so that all females contribute to the next generation with no reduction in fertility. The recursion relations connecting genotype frequencies over successive generations are

$$
\begin{aligned}
u' &= \alpha u + \tfrac{1}{4}\beta v + (1-\alpha)u(u+\tfrac{1}{2}v) + (1-\beta)\tfrac{1}{2}v(u+\tfrac{1}{2}v),\\
v' &= \tfrac{1}{2}\beta v + (1-\alpha)u(w+\tfrac{1}{2}v) + (1-\beta)\tfrac{1}{2}v + (1-\gamma)w(u+\tfrac{1}{2}v),\\
w' &= \gamma w + \tfrac{1}{4}\beta v + (1-\gamma)w(\tfrac{1}{2}v + w) + (1-\beta)\tfrac{1}{2}v(w+\tfrac{1}{2}v).
\end{aligned}
\tag{3.12}
$$

Some algebraic manipulations reveal that there exists at most one nontrivial equilibrium given by

$$
\begin{aligned}
\hat{u} &= \frac{(L+\gamma-\alpha)(\gamma-\beta)(2-\alpha-\beta)}{L[L(4-\alpha-\gamma)-(\gamma-\alpha)^2]}, \quad \hat{v} = \frac{(L+\gamma-\alpha)(L+\alpha-\gamma)(2-\alpha-\gamma)}{L[L(4-\alpha-\gamma)-(\gamma-\alpha)^2]},\\
\hat{w} &= \frac{(L+\alpha-\gamma)(\alpha-\beta)(2-\gamma-\beta)}{L[L(4-\alpha-\gamma)-(\gamma-\alpha)^2]},
\end{aligned}
\tag{3.13}
$$

where $L = (1-\alpha)(\gamma-\beta) + (1-\gamma)(\alpha-\beta)$.

The equilibrium (3.13) exists and is globally stable if $L + \gamma - \alpha < 0$ and $L + \alpha - \gamma < 0$ hold.

The symmetrical case $\alpha = \gamma$ is especially interesting. For $\alpha = \gamma < \beta$ the equilibrium simplifies to

$$
\hat{u} = \hat{w} = \frac{1}{2(2-\alpha)}, \qquad \hat{v} = \frac{1-\alpha}{2-\alpha}.
$$

which is independent of the parameter β and is stable. The symmetric multi allele version of this model can also be analyzed.

IV. INCOMPATIBILITY SYSTEMS AND SELF STERILITY

When not all possible matings can take place, incompatibility mechanisms usually operate for the prohibition of certain matings. An example which springs to mind is the human population where male-male and female-female incompatibility are in force and only male-female matings can occur. There are many other subtle incompatibilities in nature, especially involving plant populations (e.g., see East [**8**]), and we now study some simple mathematics of this phenomena.

1. A pollen elimination model. Consider a plant species in which the phenotype

in question is controlled by a single diploid locus at which there are three possible alleles A, B and C. Each plant produces both pollen and ova, but we prohibit the mating between a given pollen grain and an ovule of a plant whose genotype contains the same allele as the pollen. The model decrees that an ovule of a plant of type AB may be fertilized by only pollen of type C so that the offspring will be $\frac{1}{2}AC$ and $\frac{1}{2}BC$.

Suppose now that at the nth generation we have x_n, y_n and z_n as the proportions of AB, AC and BC respectively and suppose further that *all ova are fertilized.* It is trivial to verify that

$$x_{n+1} = \frac{y_n}{2} + \frac{z_n}{2} = \frac{1 - x_n}{2} = -\tfrac{1}{2}x_n + \tfrac{1}{2}. \tag{4.1}$$

Iterating and by symmetry we obtain

$$\begin{aligned} x_n &= \tfrac{1}{3} + (x_0 - \tfrac{1}{3})(-\tfrac{1}{2})^n, \quad y_n = \tfrac{1}{3} + (y_0 - \tfrac{1}{3})(-\tfrac{1}{2})^n, \\ z_n &= \tfrac{1}{3} + (z_0 - \tfrac{1}{3})(-\tfrac{1}{2})^n. \end{aligned} \tag{4.2}$$

Hence x_n, y_n and z_n all converge to $\frac{1}{3}$, at an oscillating geometric rate.

So far the incompatibility we have discussed arises as an incompatibility between the diploid genotype of the ovule and the haploid genotype of the pollen. Thus pollen of the incompatible type, although it contacts the female organ of the plant, dies leaving the ova intact and still available for a compatible fertilization. The incompatibility is determined by the genotype of the diploid ovule. The type of incompatibility system described above occurs in the tobacco plant (*nicotiana*).

2. A zygote elimination model. We next examine the case in which the chance of a mating is proportional to the product of the relative frequencies of both parents subject to the same incompatibility as before. In this case the chance that an AB female mates with the male genotypes AC or BC is proportional to $x(y + z) = x(1 - x)$. Table 7 is relevant at the nth generation.

Females		Frequencies of mating	Offspring	
x_n	AB	$x_n(1 - x_n)$	$\frac{AC}{2}$,	$\frac{BC}{2}$
y_n	AC	$y_n(1 - y_n)$	$\frac{AB}{2}$,	$\frac{BC}{2}$
z_n	BC	$z_n(1 - z_n)$	$\frac{AB}{2}$,	$\frac{AC}{2}$

TABLE 7

From Table 7 we find the frequencies in the next generation:

$$Nx' = \tfrac{1}{2}y(1-y) + \tfrac{1}{2}z(1-z), \quad Ny' = \tfrac{1}{2}x(1-x) + \tfrac{1}{2}z(1-z),$$
$$Nz' = \tfrac{1}{2}x(1-x) + \tfrac{1}{2}y(1-y), \tag{4.3}$$

where N is the normalizing constant $1 - x^2 - y^2 - z^2$ measuring the loss in fertility due to the diploid-diploid incompatibility.

Subtracting the pairs of equations readily shows that if $y > x$ then $x' > y'$ in the next generation and similarly if $z > x$ then $x' > z'$, etc.

Suppose, for definiteness that $z_0 < y_0 < x_0$ in the initial generation and so, $\min(x_n, z_n) < y_n < \max(x_n, z_n)$ in every succeeding generation. Clearly $y_0 \leqq \frac{1}{2}$ and therefore

$$\frac{y_0}{2(1 - x_0^2 - y_0^2 - z_0^2)} \leqq \frac{1}{4(1 - y_0)} \leqq \frac{1}{2} \tag{4.4}$$

we deduce that $|x_1 - z_1| \leqq \frac{1}{2}|x_0 - z_0|$, and so

$$|x_{n+1} - z_{n+1}| \leqq \frac{1}{2}(x_n - z_n) \leqq \frac{1}{2^n}|x_0 - z_0|$$

which implies that $x_n \to \frac{1}{3}$, $z_n \to \frac{1}{3}$ and $y_n \to \frac{1}{3}$ at a geometric rate.

Model 1 is an example of what is called **pollen elimination** since unsuitable pollen is not accepted while the ova remains intact until compatible pollen arrives. Model 2 corresponds to that called **zygote elimination** as pollen derived from an incompatible parent destroys the contacted ova.

3. A multi allelic self sterility model. In practice the number of alleles in a self sterility system of the kind discussed in IV §1 is much larger than 3. In fact as many as 35 alleles have been identified in a sample of 500 plants of *Oxalis Rosa*. We now consider a multi-allele version of IV §1 where once again it is assumed that all ova are fertilized.

Let the r alleles be denoted by $A_1, A_2, \cdots, A_r$. Then our model postulates that an $A_1 A_2$ ovule may be fertilized by $A_3, A_4, \cdots, A_r$ pollen only, etc. Let s_{ij} be the frequency of the A_iA_j genotype and we distinguish between A_iA_j and A_jA_i. Hence $s_{ii} = 0$, $\sum_{ij} s_{ij} = 2$. Then, at a given generation, the frequency of the pollen containing A_i is $q_i = \frac{1}{2}\{\frac{1}{2}\sum_j s_{ij} + \frac{1}{2}\sum_j s_{ji}\}$. Now noting that $\sum_j s_{ij} = \sum_j s_{ji}$ we have

$$q_i = \frac{1}{2}\sum_j s_{ij}. \tag{4.5}$$

We next calculate the frequency s_{ij} of the A_iA_j genotype in the next generation. The frequency of a particular ovule, say A_iA_k, in the present generation is s_{ik}. This ovule will produce one half A_i gametes and one half A_k gametes. The proportion of A_j pollen which is available to the ovule is taken to be the probability of its being

fertilized by A_j pollen. Since the proportion of compatible pollen is $1 - q_i - q_k$ we have $q_j/(1 - q_i - q_k)$ for the frequency of compatible pollen which will produce the desired A_iA_j zygote. Thus, from the A_iA_k ovule we expect a frequency $\{s_{ik}q_j/(1 - q_i - q_k)\}\frac{1}{2}$ of A_iA_j zygotes. Note that A_jA_k, A_kA_i, A_kA_j ovules also produce A_iA_j zygotes. Combining and simplifying we obtain (when $i \neq j$) the recursion relations

$$(4.6)\quad s'_{ij} = \frac{1}{2} \sum_{k \neq i,j} s_{ik} \frac{q_j}{1 - q_i - q_k} + \frac{1}{2} \sum_{k \neq i,j} s_{jk} \frac{q_i}{1 - q_j - q_k}, \qquad i,j \in 1, \cdots, r.$$

The following facts can be checked directly. For any $(l \leqq r)$,

$$s_{ij} = \frac{2}{l(l-1)} \qquad \text{for } i \neq j,\ s_{ii} = 0$$

is a fixed point of (4.6) where the indices i,j vary over a subset l of the original indices and the other frequencies are zero.

It can be shown that the gene frequency $q_i = 1/r$ $(i = 1, 2, \cdots, r)$, $r \geqq 3$ is a locally stable equilibrium. The problem of global stability has not been settled as yet.

4. Sex Determination Models. The first mathematical analysis of diploid-diploid incompatibility systems were concerned with certain naturally occurring plant genetic systems (see Fisher [**13**], Finney [**12**], Bodmer [**1**]). Subsequent investigators treated such models as special cases of the more general phenomenon of negative assortative matings, the most prominent being that of the XX, XY determination of sex in humans; although this is undoubtedly the most familiar diploid-diploid incompatibility many organisms exhibit other forms of sex determination and associated incompatibility mechanisms. The genotypes can be considered to be partitioned into two sets, with matings possible only between individuals in different sets although at random within this restriction. In the terminology set previously the models are of the zygote elimination type. For a biological justification of this formulation, see Scudo [**29**].

The first model treated here is extremely simple. We allow three genotypes AA, AB and BB, but the only matings producing viable offspring are those between a homozygote and heterozygote.

MODEL Γ

Set 1		Set 2
AA	BB	AB

TABLE 8

Matings are possible only between members of different sets. If the frequencies of the AA, AB and BB in the nth generation are respectively u_n, v_n, w_n, we obtain the recursion relations

$$\begin{aligned} T_{n-1}u_n &= u_{n-1}v_{n-1}, \\ T_{n-1}v_n &= u_{n-1}v_{n-1} + w_{n-1}v_{n-1}, \qquad (4.7) \\ T_{n-1}w_n &= w_{n-1}v_{n-1}, \end{aligned}$$

where T_{n-1} is a normalizing constant inserted to keep everything in terms of frequencies.

Obviously $u_n/w_n = u_0/w_0 = \alpha$ and $v_n = \frac{1}{2}$ for $n \geqq 1$. It follows that

$$u_n = \frac{\alpha}{2(1+\alpha)}, \quad w_n = \frac{1}{2(1+\alpha)} \text{ for } n \geqq 1.$$

Significant changes occur when a third allele is incorporated into the above model. We consider two cases according to whether the third allele C is introduced into set 1 (model Γ_1) or set 2 (model Γ_2). In the model Γ_1 the incompatibility is specified by Table 9.

MODEL Γ_1

	set 1					set 2
genotype	AA	BB	BC	AC	CC	AB
nth generation frequency	u_n	w_n	x_n	y_n	z_n	v_n

TABLE 9

Again matings are considered to take place only between individuals in different sets. The relations connecting genotype frequencies over successive generations are

$$T_{n-1}v_n = v_{n-1}u_{n-1} + v_{n-1}w_{n-1} + \frac{v_{n-1}(x_{n-1} + y_{n-1})}{2},$$

$$T_{n-1}u_n = v_{n-1}u_{n-1} + \frac{v_{n-1}y_{n-1}}{2}, \; T_{n-1}w_n = v_{n-1}w_{n-1} + \frac{v_{n-1}x_{n-1}}{2}, \qquad (4.8)$$

$$T_{n-1}x_n = \frac{v_{n-1}x_{n-1} + v_{n-1}y_{n-1}}{2}, \; T_{n-1}y_n = \frac{v_{n-1}x_{n-1} + v_{n-1}y_{n-1}}{2},$$

$z_n = 0$ for $n > 1$, where $T_{n-1} = 2v_{n-1}(1 - v_{n-1})$ is the normalizing factor. Notice that $u_n + w_n = v_n$ and $x_n = y_n$ for $n \geqq 1$. Hence

$$\frac{u_n}{x_n} = \frac{u_{n-1}}{x_{n-1}} + \frac{1}{2} = \frac{u_{n-2}}{x_{n-2}} + 1 = \cdots = \frac{n}{2} + \frac{u_0}{x_0} \text{ and } \frac{w_n}{x_n} = \frac{n}{2} + \frac{w_0}{x_0}.$$

Therefore $x_n \to 0$ and then $y_n \to 0$, so that $u_n + v_n + w_n = 2v_n \to 1$ or $v_n \to \frac{1}{2}$. Since

$$\frac{u_n}{w_n} = \frac{\dfrac{u_0}{x_0} + \dfrac{n}{2}}{\dfrac{w_0}{x_0} + \dfrac{n}{2}} \to 1 \text{ as } n \to \infty$$

we have $u_n \to \frac{1}{4}$, $w_n \to \frac{1}{4}$. The ultimate configuration of the population is therefoer $u_e = w_e = \frac{1}{4}$, $v_e = \frac{1}{2}$ and is independent of the initial makeup of the population.

Note that the previous continuum of fixed points of model Γ is reduced to the single point $u_e = w_e = \frac{1}{4}$, $v_e = \frac{1}{2}$. In model Γ the equilibrium point (which depends on the initial conditions) is achieved in one generation. In model Γ_1 the third allele disappears quite slowly at an algebraic rate.

The incorporation of the third allele into model Γ to form model Γ_1 profoundly alters the equilibrium behavior, as shown above. The only change from model Γ_1 in constructing model Γ_2 is the set to which C has been added. There are three families of equilibrium points, and the initial conditions determine which is reached. The equilibrium behavior differs markedly from that of the previous model. The following is a brief discussion of the results obtained.

MODEL Γ_2

	Set 1		Set 2			
genotype	AA	BB	AB	AC	BC	CC
nth generation frequencies	u_n	w_n	v_n	y_n	x_n	z_n

TABLE 10

There exist exactly three families of equilibria

$$F_1: \hat{w} = \tfrac{1}{2},\ \hat{v} + \hat{x} = \tfrac{1}{2}, \qquad F_2: \hat{u} = \tfrac{1}{2},\ \hat{v} + \hat{y} = \tfrac{1}{2},$$

$$F_3: \hat{u} + \hat{w} = \tfrac{1}{2},\ \hat{v} = \tfrac{1}{2}$$

and the vector $(u_n, w_n, v_n, x_n, y_n)$ always converges as $n \to \infty$. It is possible to determine precise domains of attraction to the respective equilibria.

In fact if $u_0/w_0 \leqq 1$, $y_0/x_0 \leqq 1$ and $u_0y_0/w_0x_0 < 1$, then the limiting equilibrium is in F_1. Symmetrically, if $w_0/u_0 \leqq 1$, $x_0/y_0 \leqq 1$ and $w_0x_0/u_0y_0 < 1$ then the limiting equilibrium belongs to F_2. If $(u_{n-1}/w_{n-1}) - 1$ and $(u_n/w_n) - 1$ alternate continually as $n \to \infty$ or are zero, then the limit equilibrium belongs to F_3 (see Karlin **[16]** for proofs). The domain of attraction to F_3 is usually a hypersurface.

One final example is where each sex is characterized by three genotypes as follows. (A nine allele expression of this model arises in a strain of wasp. For certain fungi, including yeast, sex determination appears to be controlled at a single locus.)

	Set 1			Set 2		
	AA	BB	CC	AB	AC	BC
frequency	x	y	z	u	v	w

The recurrence relations are as follows:

$$\begin{aligned} Tx' &= x(u+v), & Tu' &= (x+y)\,u + vy + wx, \\ Ty' &= y(u+w), & Tv' &= (x+z)\,v + uz + xw, \\ Tz' &= z(v+w), & Tw' &= (y+z)\,w + vy + uz, \end{aligned}$$

$$T = 2(x+y+z)\,(u+v+w).$$

The stable equilibria are precisely the fixed points

$$\begin{aligned} x+y &= \tfrac{1}{2} & u &= \tfrac{1}{2}; & x &= \tfrac{1}{2} & v+w &= \tfrac{1}{2} \\ x+z &= \tfrac{1}{2} & v &= \tfrac{1}{2}; & y &= \tfrac{1}{2} & u+w &= \tfrac{1}{2} \\ y+z &= \tfrac{1}{2} & w &= \tfrac{1}{2}; & z &= \tfrac{1}{2} & v+w &= \tfrac{1}{2}. \end{aligned}$$

An interior unstable fixed point $x = y = z = 1/9$, $u = v = w = 2/9$, also exists.

It can be proved generally in the case of three alleles at a single locus that, any grouping for sex determination exhibits only the $\frac{1}{2}$ sex ratio in a stable configuration. When sex is determined involving at least two loci, then a stable sex ratio may be different from $\frac{1}{2}$.

V. MUTATION SELECTION BALANCE

1. Mutation balance. We assume that each A allele has a probability μ of mutating to B (and hence $1-\mu$ of not mutating), that ν similarly is the mutation rate of B to A and that no other forces are acting to change gene frequencies. It is easily seen that

$$p_n = (1-\mu)\,p_{n-1} + \nu(1-p_{n-1}), \tag{5.1}$$

where p_n is the gene frequency of A in the nth generation. This equation can be rewritten in the form

$$\left(p_n - \frac{\nu}{\mu+\nu}\right) = (1-\mu-\nu)\left(p_{n-1} - \frac{\nu}{\mu+\nu}\right) = (1-\mu-\nu)^n\left(p_0 - \frac{\nu}{\mu+\nu}\right).$$

Thus $p_n \to \nu/(\mu+\nu)$ as $n \to \infty$ at a rate $(1-(\mu+\nu))^n$ i.e., $p_n - 1/(\mu+\nu)$ is of order $(1-\mu-\nu)^n$. There is thus a stable intermediate equilibrium point, whose position depends on the ratios of the two mutation rates. However, since mutation rates are generally less than 10^{-5}, the rate of convergence to the equilibrium is exceedingly slow. As we shall see below, it seems likely that selection differentials are nearly always large enough to mask these balancing effects of opposing mutation rates.

2. Immigration balance. We assume that a proportion m of the population is replaced in each generation by individuals from another population with constant A and B gene frequencies P and Q respectively. The change in gene frequency is then given by

$$p_n = (1 - m)\, p_{n-1} + mP. \tag{5.2}$$

As $n \to \infty$, then $p_n \to P$, the frequency of the immigrant population, at a rate $(1 - m)^n$. If we put $v = mP$ and $\mu = m(1 - P)$ then equation (5.2) is identical to equation (5.1), so that this situation is exactly analogous to the mutation balance. Both factors cause linear changes in the gene frequencies.

3. Mutation-selection balance for disadvantageous genes. Assume genotypes AA, AB and BB have relative fitnesses 1, $1 - hs$, and $1 - s$ where $s, h \geqq 0$, and that p and q are the gene frequencies in fertilized zygotes. Gene frequencies are measured in the gametes which combine at random to form the fertilized zygote, before selection has acted, and mutation is assumed to occur after selection during the formation of the next generation's gametes.

As in Section 2, the gene frequencies of A and B after selection, before mutation, are

$$\frac{p^2 + (1 - hs)pq}{1 - 2hspq - sq^2} \text{ and } \frac{(1 - s)q^2 + (1 - hs)pq}{1 - 2hspq - sq^2},$$

respectively. Allowing only one way mutation $A \to B$, the new frequency of B will be

$$q' = \frac{(1 - s)q^2 + (1 - hs)pq}{1 - 2hspq - sq^2} + \mu\frac{[p^2 + (1 - hs)pq]}{1 - 2hspq - sq^2}. \tag{5.2}$$

Equilibria are obtained as the solutions of

$$sq^3(2h - 1) + sq^2[1 - 3h - h\mu] + q[\mu + hs(1 + \mu)] - \mu = 0. \tag{5.3}$$

The mutation rate μ is always very small. One stable equilibrium is approximately

$$q \sim \mu/hs \tag{5.4}$$

provided μ is small compared with hs. The nth generation frequency q_n approaches its equilibrium value of μ/sh at a geometric rate of approximate order 1–sh. It is noteworthy that this solution depends only on the product sh and not on s alone, indicating that the fitness of the heterozygote dominates the situation. Given q and hs for any particular gene, assumed to have reached its equilibrium frequency, we can estimate from (5.4) the magnitude of the mutation rate μ. This was, in fact, the way that mutation rates in man were originally derived by Danforth in 1920 and later by Haldane.

When $h = 0$ the allele B is recessive with respect to its effect on fitness and (5.4) reduces to $(q - 1)(sq^2 - \mu) = 0$. The solution $q = \sqrt{\mu/s}$ is the only stable equilibrium, of course provided $\mu \leqq s$.

The results of this model have been frequently applied in estimating the mutation rate for recessive human diseases.

Criteria for selection mutation balance for a character controlled at two loci are given in Karlin and McGregor [**21**]. In Section 8 we present a model for mutation selection balance involving an infinite number of types. Those considerations are also relevant to an understanding of polygenic inheritance (characters determined by many loci).

VI. THE CONCEPT OF IDENTITY BY DESCENT AND APPLICATIONS

The **inbreeding coefficient** of an individual (introduced first by Wright) is defined to be the probability that two genes at a single locus are **identical by descent** by which we mean that the genes can be traced back to copies of the same gene in a particular individual of a previous generation. Certain finite size population genetic problems can be solved relatively easily using calculations for probabilities of descent. We expose a series of important models exemplifying the method. (This method has been exploited by many including Malécot, Kimura, Kempthorne and others. See Karlin [**16**] and [**17**] for further applications and references on this subject.)

1. Monoecious diploid finite population. A **monoecious individual** is one that can contribute both male and female gametes (e.g., as occurs commonly in plants).

Consider a population of N monoecious individuals diploid at an autosomal locus, reproducing randomly but maintaining constant population size. More specifically we may stipulate that each individual produces an infinite number of copies of each of his genes to form a pool from which the next generation is formed by choosing N pairs at random where each parental gene is represented to the extent of $\frac{1}{2}N^{-1}$-th of the complete gene pool.

Let I_t denote the probability that two homologous chromosomes at a given locus in an individual in the tth generation carry genes identical by descent. Let J_t be the probability that two homologous chromosomes of the tth generation, chosen at random one from each of two different individuals, carry genes identical by descent.

Under random mating two genes are derived from the same parental individual with probability $1/N$ or from different individuals with probability $1 - 1/N$. In the former event either they are copies of the same gene or they are copies of the homologous pair, each occurring with probability $\frac{1}{2}$. We may evidently compute I_t and J_t according to the same recursion relations

$$J_t \text{ and } I_t = \frac{1}{N}\left(\frac{1}{2} + \frac{1}{2}I_{t-1}\right) + \left(1 - \frac{1}{N}\right) J_{t-1}, \quad t \geqq 1. \tag{6.1}$$

Thus $I_t = J_t$ for $t \geqq 1$ and (6.1) reduces to

$$I_t = \tfrac{1}{2}N^{-1} + (1 - \tfrac{1}{2}N^{-1})I_{t-1}, \qquad t \geqq 2. \tag{6.2}$$

We introduce the quantity $H_t = 1 - I_t$ and then (6.2) is converted into

$$H_t = (1 - \tfrac{1}{2}N^{-1})H_{t-1} = (1 - \tfrac{1}{2}N^{-1})^{t-1}H_1, \qquad t \geqq 1, \tag{6.3}$$

where $H_1 = 1 - I_1$ and $I_1 = \frac{1}{2}N^{-1}(1 + I_0) + (1 - N^{-1})J_0$. Equation (6.3) shows that H_t tends to zero at a geometric rate $(1 - \frac{1}{2}N^{-1})$.

The above analysis implies two interesting conclusions. Firstly, the ultimate population is composed exclusively of inbred individuals, i.e., individuals with inbreeding coefficient 1. Secondly, even for the process of random mating, limitation of population size imposes a certain degree of inbreeding which eliminates, at an exponential rate, the heterozygote types.

2. Dioecious finite diploid population. We consider a two sex population consisting of N_1 males and N_2 females. Let I_t be the probability that two homologous genes from the same male or female of the tth generation are identical by descent. Let J_t be the probability that two genes chosen at random one from each of two different males or females in the tth generation are identical by descent. Let K_t be the probability that two genes chosen at random, one from a male, the other from a female of the tth generation, are identical by descent. Finally, let $\tilde{J}_t$ denote the probability that two genes chosen at random in the tth generation, one each from different individuals (with no reference to sex), are identical by descent. Symmetry suggests and indeed it can be easily proved that I_t and $J_t = \tilde{J}_t$ are well defined.

We now develop recursion formulas for the quantities introduced above by examining the source of the two genes in a given individual traced two generations back. Consider two genes in a given individual. Conditional that they both come from males, two generations back, the probability they derive from the same male (say A) is $(N_1/N_1^2) = N_1^{-1}$.

The probability is $\frac{1}{4}$ that two children B and C of A transmit to their offspring D the genes received from A. Now the genes given B and C by A are copies of the same gene or correspond to distinct homologous genes with probability $\frac{1}{2}$ each. In the latter event the genes are identical by descent with probability I_{t-2}. This accounts for the first term of the recursion relation

$$I_t = \tfrac{1}{4}N_1^{-1}(\tfrac{1}{2} + \tfrac{1}{2}I_{t-2}) + \tfrac{1}{4}N_2^{-1}(\tfrac{1}{2} + \tfrac{1}{2}I_{t-2}) + (1 - \tfrac{1}{4}N_1^{-1} - \tfrac{1}{4}N_2^{-1})J_{t-2}. \tag{6.4}$$

The second term reflects the circumstance when both genes derive from the same female parent. The probability is $(1 - \frac{1}{4}N_1^{-1} - \frac{1}{4}N_2^{-1})$ that the two genes of D derive from distinct individuals of the $(t-2)$-th generation, in which case the probability is J_{t-2} that they are identical by descent.

A similar kind of reasoning establishes the relation

$$J_t = \tfrac{1}{4}N_1^{-1}(\tfrac{1}{2} + \tfrac{1}{2}I_{t-1}) + \tfrac{1}{4}N_2^{-1}(\tfrac{1}{2} + \tfrac{1}{2}I_{t-1}) + (1 - \tfrac{1}{4}N_1^{-1} - \tfrac{1}{4}N_2^{-1})J_{t-1}. \tag{6.5}$$

Notice that the subscript on the right now involves the $(t-1)$-th generation rather than the $(t-2)$-th.

The identical formula as in (6.5) obtains with the left side replaced by K_t. It follows that $J_t = K_t$. Comparing (6.5) and (6.4) we may conclude that $J_{t-1} = I_t$ and then we rewrite (6.4) in the form

$$I_t = N_e^{-1}(\tfrac{1}{2} + \tfrac{1}{2}I_{t-2}) + (1 - N_e^{-1})I_{t-1}, \tag{6.6}$$

where $N_e^{-1} = \frac{1}{4}N_1^{-1} + \frac{1}{4}N_2^{-1}$, a quantity commonly called the **effective population number**. Let $H_t = 1 - I_t$ and then (6.6) becomes

$$H_t = (1 - N_e^{-1})H_{t-1} + \tfrac{1}{2}N_e^{-1}H_{t-2}, \qquad t \geqq 2. \tag{6.7}$$

The solution of this second order difference equation has the form $H_t = a\lambda_1^t + b\lambda_2^t$, $t \geqq 2$, where λ_i $(i = 1,2)$ are roots of the quadratic equation $\lambda^2 - (1 - N_e^{-1})\lambda - \frac{1}{2}N_e^{-1} = 0$. Hence as $t \to \infty$, H_t behaves asymptotically as

$$H_t \sim \tfrac{1}{2}a[1 - N_e^{-1} + (1 + N_e^{-2})]^t. \tag{6.8}$$

The special case of sib mating arises when $N_1 = N_2 = 1$ and so $N_e = 2$. Then $H_t \sim a(\frac{1}{4}(1 + \sqrt{5}))^t$.

3. Loss of k alleles out of p in a haploid model. Consider a finite constant size (say N individuals) **haploid** population (each individual carries one dose of an allele) undergoing some general pattern of reproduction where the number of alternative alleles represented in the population is at least $p > 2$. We investigate the problem of determining the rate at which k of the p alleles are lost from the population.

The reproduction mechanism is as follows. Each individual replicates his type in some general fashion but with no selection differences operating among the types. The next generation is formed by choosing at random N progeny from the output of the previous generation. The parameters of the reproduction mechanism are the numbers g_{ij} = to the probability that i randomly chosen progeny derive from j distinct parents $(i,j = 1,2,\cdots,N)$. Obviously $g_{ij} = 0$ for $j > i$ so the matrix $G = \| g_{ij} \|_1^N$ is lower triangular. Clearly $g_{11} = 1$ and we postulate that

$$g_{kk} > g_{k+1,k+1} > 0 \qquad (k = 1,2,\cdots,N-1) \text{ and } g_{k,k-1} > 0 \tag{6.9}$$

in order to avoid pathological algebraic annoyances. These conditions are satisfied in almost all examples. In the special case where each parent contributes exactly r replicas of his own type then an elementary combinatorial analysis shows that

$$g_{ij} = \begin{cases} \Sigma^* \dfrac{\binom{r}{i_1}\binom{r}{i_2}\cdots\binom{r}{i_j}\binom{N}{j}}{\binom{Nr}{i}} & i \geqq j \\ 0 & i < j, \end{cases} \tag{6.10}$$

where Σ^* indicates summation over all $i_1, i_2, \cdots, i_j \geqq 1$ subject to $i_1 + i_2 + \cdots + i_j = i$.

The conditions of (6.9) are obviously satisfied in this circumstance. Notice that here $g_{ii} \to (N(N-1)\cdots(N-i+1))/N^i$ as $r \to \infty$.

Let $P_{ij}^{(t)}$ be the probability that i randomly chosen different individuals of the tth generation consist of j different types (alleles). Our *objective is to ascertain the asymptotic properties* of $P_{Nj}^{(t)}$ as $t \to \infty$ for $j = 1, 2, \cdots, p$. Since the population size is kept constant we expect ultimate fixation in one type, i.e., $P_{Nj}^{(t)} \to 0$ as $t \to \infty$ for $j = 2, \cdots, N$. We wish to determine the rate of this approach to zero. The key to the analysis is the recursion relation

$$(6.11) \qquad P_{ij}^{(t+1)} = \sum_{k=1}^{N} g_{ik} P_{kj}^{(t)} \qquad (i, j, \cdots, N).$$

The derivation is simple and follows by considering the various possibilities describing the parental genes that can produce the given sampled genes.

If we introduce the matrices $P^{(t)} = \| P_{ij}^{(t)} \|$, then (6.11) can be written concisely as the matrix product $P^{(t+1)} = GP^{(t)}$, and iteration produces

$$(6.12) \qquad P^{(t)} = G^t P^{(0)},$$

where G^t is the tth power of the matrix G and $P^{(0)}$ provides the information of the initial frequencies of types. Since G is lower triangular and the diagonal elements are distinct by assumption, we may conclude that the eigenvalues of G are $\lambda_1 = g_{11} = 1$, $\lambda_2 = g_{22}, \cdots, \lambda_k = g_{kk}, \cdots, \lambda_N = g_{NN}$.

A system of left eigenvectors of G can be constructed of the form

$$v^k = (v_1^{(k)}, \cdots, v_k^{(k)}, 0, \cdots, 0), \qquad k = 1, 2, \cdots, N$$

with the property $v_k^{(k)} \neq 0$. This last fact derives from the condition $g_{kk} > g_{k-1,k-1}$. Let V be the matrix with row vectors $v^{(1)}, v^{(2)}, \cdots, v^{(N)}$ and $U = V^{-1}$. Since V is lower triangular, so is U. Of course, $G = U \Omega V$, where Ω is the diagonal matrix of eigenvalues of G whose values are $g_{11}, g_{22}, \cdots, g_{NN}$. It is not difficult to prove inductively that $u_i^{(k)} > 0$ for all $i \geqq k$. Consider now

$$P_{Nj}^t = \sum_{k=1}^{N} G_{Nk}^{(t)} P_{kj}^0 = \sum_{k=j}^{N} G_{Nk}^{(t} P_{kj}^0.$$

Expanding

$$\begin{aligned} G_{Nj}^t &= \sum_{k=1}^{N} u_N^k g_{kk}^t v_k^{(j)} = \sum_{k=j}^{N} u_N^{(k)} [g_{kk}]^t v_k^{(j)} \\ &= [g_{jj}^t] u_N^{(j)} + O[g_{j+1,j+1}]^t \end{aligned}$$

where the last reduction is valid since $v_k^{(j)} = 0$ for $k < j$. Since $u_N^{(j)} > 0$ we have proved the following theorem.

THEOREM. *Suppose* (6.9) *holds. If* $P_{jj}^0 > 0$ *then the probability that a population*

of N haploid individuals contains at least j types in the t-th generation is of the order of magnitude $c_j[g_{jj}]^t$ *where* c_j *is a positive constant depending on the initial set of frequencies.*

The condition $P_{jj}^0 > 0$ for $j \leqq p$ is very weak and would ordinarily be satisfied. For further discussion of this model and ramifications we refer to Karlin [**17**] Section 6, and Felsenstein [**11**].

5. Identity by descent and mutation effects. Consider a population of N diploid individuals or $2N$ genes with an infinite series $A_1, A_2, \cdots$ of possible alleles at a locus with no selective differences among the allelic types. The population is randomly reproducing as in Model I, i.e., the $2N$ genes of the next generation are formed by repeated sampling with replacement from the $2N$ genes of the present generation. Suppose moreover that as each gene is drawn there is a probability u that a mutation occurs and any new mutant allele is of a not previously existing type.

Let I_t be the probability in generation t that two genes sampled at random are identical by descent. A recursion formula analogous to (6.1) with due account of mutation is

$$I_t = \left[\frac{1}{N}\left(\frac{1}{2} + \frac{1}{2}I_{t-1}\right) + \left(1 - \frac{1}{N}\right)I_{t-1}\right](1-u)^2.$$

Letting $t \to \infty$, we get the equilibrium value $\lim_{t\to\infty} I_t = I$, where

$$I = \frac{(1-u)^2}{1 + 4Nu - 2Nu^2}$$

and for u small and N large such that $4Nu = \theta$ we have the approximate formula $I = 1/(1+\theta)$.

Of considerable interest for discussions relevant to non-Darwinian evolution (Neutral mutation theory) is the evaluation of the probability

$$(6.13) \qquad P\{2N, u, n_1, n_2, \cdots, n_k\}$$

that a sample of r genes, chosen from the population, contains just k different allelic types with n_1 of one kind, n_2 of a second kind and so on, n_k of a kth kind where the n_i are positive integers with sum r. For the significance of the computation of (6.13) and its utility in evaluating the relevance of neutral mutation theory, we refer to Ewens [**10**]. The quantity (6.13) is a complicated function of $2N$ and u. However, if we let $N \to \infty$ and $u \to 0$ in such a way that $4Nu$ converges to a finite non-zero limit θ, then (6.13) converges to a relatively simple limit formula

$$(6.14) \qquad P(\theta; n_1, n_2, \cdots, n_k) = \frac{r!}{n_1 n_2 \cdots n_k\, \alpha_1!\, \alpha_2! \cdots \alpha_p!}\frac{\theta^k}{L_r(\theta)},$$

where p is the number of distinct integers in the set $\{n_1, n_2, \cdots, n_k\}$ of which there are

exactly α_1 indices equal to an integer, α_2 indices equal to a different integer, and so on, and exactly α_p indices equal to the pth distinct value among the numbers $n_1, n_2, \cdots, n_k$. Here

$$L_r(\theta) = \theta(\theta + 1)(\theta + 2)\cdots(\theta + r - 1).$$

The formula was suggested by Ewens [**10**] and rigorously established in Karlin and McGregor [**22**]. The method relies heavily on the concept of identity by descent.

VII. EVOLUTION OF A POPULATION WITH POLYGENIC CHARACTERS

1. A model of a polygenic trait. Consider a population with an infinite number of possible types. Assume that the different types are identified with points of the real line $\mathbb{R}$. One example is where the type x can be associated with the "fitness" of the given individual. A second case is where x corresponds to a measurable numerical trait whose value is determined by the combined effects of many loci.

Consider the frequency distribution of the types in the population. More precisely, let A be any interval (or Borel measurable set) in $\mathbb{R}$ and let $m_t(A)$ be the proportion of the population (population size is for our purposes, regarded of large-infinite size) of types corresponding to A at generation t. Selection and mutation affect changes in m_t over successive generations in the following manner:

(i) The relative viability of an offspring of type x compared to that of type y is in the ratio $\gamma(x)/\gamma(y)$ which we stipulated as a first approximation to be independent of time. Assuming each parental type replicates its identical type, the change of frequency distribution due to this selection is to be calculated by the formula

$$\tilde{m}_{t+1}(A) = \frac{\int_A \gamma(x) m_t(dx)}{\int_{\mathbb{R}} \gamma(x) m_t(dx)}$$

for all intervals (and sets) A.

(ii) Mutation acts after selection as follows: let $p_t(B, x)$ be the conditional probability that an offspring of an x-type parent of generation t alter its form to that of type in B. Then a parent of type x, affected by selection and mutation will produce offspring of type in A is calculated modulo a proportionality constant by the expression $\gamma(x)p_t(A, x)$. It follows that the total number of A-type offspring in generation $t + 1$ is proportional to $\int_{\mathbb{R}} \gamma(x)p_t(A, x)m_t(dx)$ which after converting to frequencies, becomes

$$m_{t+1}(A) = \frac{\int_{\mathbb{R}} \gamma(x) p_t(A, x) m_t(dx)}{\int_{\mathbb{R}} \gamma(x) m_t(dx)}. \tag{7.1}$$

The evolution of the frequency distributions m_t over time is the primary object under investigation. To achieve qualitative results and deeper insights into the behavior of m_t as t increases we now specialize to the situation where

$$p_t(B,u) = \int_B dG(x-u) \text{ and } \gamma(x) = \lambda^x, \qquad \lambda > 1 \tag{7.2}$$

so that the difference between a parent and offspring has the same distribution $G(u)$ (called the **distribution of the mutation**) over the whole population. The reproduction rate of an x-type parent is λ^x so that a type is more advantageous with larger values. For the case of $\gamma(x) = \lambda^x$ the meaning of x is strongly correlated with the actual fitness of the x-individual.

Let $F_t(x)$ be the proportion of types $\leqq x$ in the population at time t. Manifestly, $F_t(x)$ is a distribution function of the variable x. Define $E_t = \int_{-\infty}^{\infty} x dF_t(x)$ as the average fitness and $V_t = \int_{-\infty}^{\infty} [x - E_t]^2 dF_t(x)$ as the fitness variance. Define for any distribution $H(x)$ the quantity $\bar{H} = \inf\{x \mid H(x) = 1\}$ as the largest point in the spectrum of $H(x)$. The following results were proved by Eshel [9], (see Karlin [23] for improvements and extensions).

THEOREM I. *Assume $\bar{F}_0 < \infty$ (i.e., the initial fitness distribution in the population is bounded). Suppose that $\bar{G} < \infty$, that is the maximal possible mutation change is bounded. Then*

$$\lim_{t \to \infty} (E_{t+1} - E_t) = \bar{G}. \tag{7.3}$$

The rate of evolution (= the rate of change of the average fitness in the population) approaches $\bar{G}$.

A more refined result pertains to the changes in the centered fitness distribution $F_t(x - E_t)$ as $t \to \infty$.

THEOREM II. *Under the assumptions of Theorem I $F_t(x - E_t)$ tends to a limit distribution $F(x)$ whose variance is finite.*

In particular the proportion of types compared to the mean fitness in any given region tends to a positive value. We state as a consequence of Theorem II: *If $\bar{G} = 0$ (i.e., all mutations are deleterious or neutral), then it follows that $F_t(x)$ approaches a limiting mutation selection balance with distribution of types $F(x)$ iff $G(x)$ has a positive jump at 0.*

The results cited above hinge strongly on the assumptions of (7.2). To what extent are corresponding conclusions valid for other choices of the selection functions $\gamma(x)$ not of exponential growth λ^x? Cases where $\gamma(x)$ is bell-shaped (e.g., $\gamma(x) = e^{-x^2}$ or $1/(1 + x^2)$) would be of interest in treating the evolution of quantitative traits where the optimum type has an intermediate value.

2. Another model of a polygenic trait. Another model of a polygenic trait involving a selection balance and the mating process proposed by Haldane has the following structure.

The set of all possible phenotypes are again identified with the real line. Let the proportion of the population exhibiting phenotype in an interval A in generation t be

$$m_t(A) = \int_A p_t(x)\,dx. \tag{7.4}$$

(For ease of exposition we have assumed the existence of a density p_t for the frequency measure $m_t(dx)$.) The basic assumption for this model is that the distribution of the type of the offspring depends on the type of each parent, through the conditional probability (**segregation function**) $L(x; x_1, x_2)\,dx$ equal to the probability that the offspring is of type x to $x + dx$ given the parental types are x_1 and x_2. Clearly,

$$\int_{-\infty}^{\infty} L(x; x_1, x_2)\,dx = 1.$$

In theory, L could be determined from careful analysis of breeding experiments. Assuming random union of types the density of phenotypes in the next generation would ordinarily be calculated by the formula

$$\check{p}_{t+1}(x) = \iint L(x; x_1, x_2) p_t(x_1) p_t(x_2) dx_1 dx_2 \tag{7.5}$$

before selection has acted. The action of selection is determined as in Model 1 by a function $\gamma(x)$ which is the relative survival probability for individuals of type x. Taking account of selection, the density $p_t(x)$ is altered to

$$\tilde{p}_t(x) = \frac{\gamma(x) p_t(x)}{\int_{-\infty}^{\infty} p_t(x)\gamma(x)\,dx}.$$

Subject to random mating, segregation (described by $L(x; x_1, x_2)$) and selection (measured in relative terms by $\gamma(x)$) we obtain the non-linear transformation law

$$p_{t+1}(x) = \frac{\iint p_t(x_1)p_t(x_2)\gamma(x_1)\gamma(x_2)L(x; x_1, x_2)\,dx_1 dx_2}{\left(\int_{-\infty}^{\infty} p_t(\xi)\gamma(\xi)d\xi\right)^2}.$$

For certain choices of $L(x; x_1, x_2)$ for a large class of bell-shaped functions $\gamma(x)$ we can deduce the fact that $m_t(x)$ converges to a limiting stable frequency distribution.

Other models for polygenic traits were studied by Kimura (see Crow and Kimura [7], pages 294–296, Slatkin [31], Haldane [14], among others).

VIII. SOME SELECTION MODELS FOR TWO LOCUS MARKERS

Consider a diploid population of a character determined by two loci with possible alleles A, a and B, b at the first and second locus respectively. There are therefore four types of chromosomes (or referred to as gametes):

$$AB \quad Ab \quad aB \quad ab \tag{8.1}$$

and 10 genotypes

$$\frac{AB}{AB}, \frac{AB}{Ab}, \frac{AB}{aB}, \frac{AB}{ab}, \frac{Ab}{Ab}, \frac{Ab}{aB}, \frac{Ab}{ab}, \frac{aB}{aB}, \frac{aB}{ab}, \frac{ab}{ab}$$

where the symbol AB/aB, for example, means that the alleles A and B sit on one of the chromosomes while the alleles a and B are found on the other. Let $M = \| m_{ij} \|_{i,j=1}^{4}$ denote the fitness matrix, where m_{ij} is the fitness of the genotype composed from the i and j type chromosomes.

Let x_1, x_2, x_3 and x_4 be the frequencies of the four gamete types in the order of (8.1). Assuming random union of gametes (= random mating) and recalling the nature of Mendelian segregation involving recombination frequency r (refer here back to Section I), it is easy to check Table 11.

Reading off from the table we find that the frequency x_1' of AB in the next generation is proportional to

$$\begin{aligned} x_1' &\sim x_1^2 m_{11} + 2x_1x_2 m_{12} \tfrac{1}{2} + 2x_1x_3m_{13}\tfrac{1}{2} + 2x_1x_4m_{14}(1-r)\tfrac{1}{2} + 2x_2x_3r\tfrac{1}{2} \\ &= x_1m_1 - rD, \end{aligned}$$

where $m_1 = \sum_{j=1}^{4} m_{ij}x_j$, $D = x_1x_4m_{14} - x_2x_3m_{23}$. Similar expressions result for x_2', x_3' and x_4'. The recursion relations connecting frequencies over successive generations become

$$x_i' = \frac{x_im_i + \varepsilon_i rD}{W}, \qquad i = 1,2,3,4, \tag{8.2}$$

where $\varepsilon_2 = \varepsilon_3 = -\varepsilon_1 = -\varepsilon_4 = 1$, $m_i = \sum_{j=1}^{4} m_{ij}x_j$, $W = \sum_{i,j=1}^{4} m_{ij}x_ix_j$.

1. No selection differences. The special case where $m_{ij} \equiv 1$ (no selection differences) is the most classical case treated. Then (8.2) reduces to

$$x_i' = x_i + \varepsilon_i rD, \qquad i = 1,2,3,4. \tag{8.3}$$

It is convenient to introduce the **gene frequency variables**

$$\begin{aligned} p_1 &= x_1 + x_2 = \text{(frequency of } A\text{)}, \; p_2 = x_1 + x_3 = \text{(frequency of } B\text{)}, \\ D &= x_1x_4 - x_2x_3 \text{ (linkage disequilibrium function)}. \end{aligned} \tag{8.4}$$

We can obviously recapture the gamete frequency according to

Mating type	Frequency	Viability	Segregation
$\frac{AB}{AB}$	x_1^2	m_{11}	AB
$\frac{AB}{Ab}$	$2x_1 x_2$	m_{12}	$\frac{1}{2}AB + \frac{1}{2}ab$
$\frac{AB}{aB}$	$2x_1 x_3$	m_{13}	$\frac{1}{4}AB + \frac{1}{2}aB$
$\frac{AB}{ab}$	$2x_1 x_4$	m_{14}	$(1-r)(\frac{1}{2}AB + \frac{1}{2}ab) + r(\frac{1}{2}Ab + \frac{1}{2}ab)$
$\frac{Ab}{Ab}$	x_2^2	m_{22}	Ab
$\frac{Ab}{aB}$	$2x_2 x_3$	m_{23}	$(1-r)(\frac{1}{2}Ab + \frac{1}{2}aB) + r(\frac{1}{2}AB + \frac{1}{2}ab)$
$\frac{Ab}{ab}$	$2x_2 x_4$	m_{24}	$\frac{1}{2}Ab + \frac{1}{2}ab$
$\frac{aB}{aB}$	x_3^2	m_{33}	aB
$\frac{aB}{ab}$	$2x_3 x_4$	m_{34}	$\frac{1}{2}aB + \frac{1}{2}ab$
$\frac{ab}{ab}$	x_4^2	m_{44}	ab

TABLE 11

(8.5)
$$\begin{aligned} x_1 &= p_1p_2 + D, & x_2 &= p_1(1-p_2) - D, \\ x_3 &= (1-p_1)p_2 - D, & x_4 &= (1-p_1)(1-p_2) + D. \end{aligned}$$

On the basis of (8.3) and (8.4) we obtain

(8.6)
$$p_1' = p_1, \; p_2' = p_2, \; D' = (1-r)D$$

and therefore $D^{(n)} = (1-r)^n D^{(0)} \to 0$ provided $r > 0$. Combining (8.6) with (8.5) we see that

$$\begin{aligned} x_1^{(n)} &= p_1p_2 + D^{(n)} \to p_1p_2 \text{ as } n \to \infty \\ x_2^n &\to p_1(1-p_2), \text{ etc.} \end{aligned}$$

Letting $p^0(A)$ $(p^0(B))$ denote the initial frequency of the A gene (B gene) etc. we can express the limiting frequencies in the form limit frequency of

$$
\begin{aligned}
&f^\infty(AB) = x_1^{(\infty)} = p_1^{(0)}p_2^0 = p^{(0)}(A)p^{(0)}(B) \\
(8.7)\qquad &f^\infty(Ab) = p^0(A)p^0(b), \qquad f^\infty(aB) = p^0(a)p^0(B) \\
&f^\infty(ab) = p^0(a)p^0(b)
\end{aligned}
$$

so that the two loci act in the limit independently. provided, recombination is positive.

2. Additive viabilities. This is the case where the fitness of a genotype is determined as the additive effects of the fitness contributed by each locus separately. Specifically, suppose $\sigma_1, \sigma_2, \sigma_3$ denote the relative fitnesses of AA, Aa, aa respectively and s_1, s_2, s_3 represent the relative fitnesses of BB, Bb and bb respectively. Then m_{11} the fitness of AB/AB is $\sigma_1 + s_1$, the sum of the fitnesses of AA and BB. Similarly, m_{14} for AB/ab is $\sigma_2 + s_2$ and m_{24} of Ab/ab is $\sigma_2 + s_3$, etc.

In the case of additive fitnesses and heterozygote advantage at each locus, i.e., $\sigma_2 > \max(\sigma_1, \sigma_3)$ and $s_2 > \max(s_1, s_3)$, it can be proved that the limiting gamete frequencies are

$$(8.8)\qquad \lim_{n\to\infty} x_1^{(n)} = \hat{p}_1\hat{p}_2, \quad \lim_{n\to\infty} x_2^{(n)} = \hat{p}_1(1-\hat{p}_2) \text{ etc.,}$$

where

$$\hat{p}_1 = \frac{\sigma_2 - \sigma_3}{2\sigma_2 - \sigma_1 - \sigma_3}, \quad \hat{p}_2 = \frac{s_2 - s_3}{2s_2 - s_1 - s_3}$$

valid for any initial frequency vector $(x_1^{(0)}, x_2^0, x_3^0, x_4^0)$ provided $x_1^0x_2^0x_3^0x_4^0 > 0$.

Other examples of viability arrays that can be mostly analyzed include the cases of multiplicative viabilities and the symmetric viability model (e.g., see Bodmer and Felsenstein [**3**], Kojima and Lewontin [**27**] and Karlin and Feldman [**19**]).

Work supported in part by the National Institute of Health Grant USPRS 10452–09.

References

1. W. F. Bodmer, (1960), Discrete stochastic processes in population genetics, J. Roy. Stat. Soc. B 22: 218–244.

2. W. F. Bodmer, (1965), Differential fertility in population genetic models, Genetics 51: 411–424.

3. W. F. Bodmer and J. Felsenstein, (1967), Linkage and selection: Theoretical analysis of the deterministic two locus random mating model, Genetics 57: 237–265.

4. C. Cannings, (1967), Equilibrium, convergence and stability at a sex-linked locus under natural selection, Genetics 56: 613–617.

5. C. Cannings, (1969), Unisexual selection at an autosomal locus, Genetics 62: 225–229.

6. L. Cavalli Sforza and W. Bodmer, (1971), Genetics of Human Population, Freeman, San Francisco.

7. J. F. Crow and M. Kimura, (1970), An introduction to population genetics theory, Harper & Row, New York.

8. E. M. East, (1940), The distribution of self sterility in flowering plants, Proc. Amer. Phil. Soc.

9. I. Eshel, (1971), On evolution in a population with an infinite number of types, Theor. Pop. Biol. 2: 209–236.

10. W. J. Ewens, (1972), The sampling theory of selectively neutral genes, Theor. Pop. Biol., March 1972.

11. J. Felsenstein, (1971), On the loss of alleles from a Haploid population, Theor. Pop. Biol., December 1971.

12. D. J. Finney, (1952), The equilibrium of a self-incompatible polymorphic species, Genetica 26: 33–64.

13. R. A. Fisher, (1941), The theoretical consequence of polyploid inheritance for the mid style form of lythrum salicaria, Ann. Eugen. 11: 31–38.

14. J. B. S. Haldane, (1924), A mathematical theory of natural and artificial selection, Biol. Proc. Camb. Phil. Soc., Biol. Sci. 1: 158–163.

15. J. B. S. Haldane, (1932), The causes of evolution, Harper & Row, New York.

16. S. Karlin, (1968), Equilibrium behavior of population genetic models with non-random mating, Part I: Preliminaries and special mating systems, J. Appl. Prob. 5: 231–313.

17. S. Karlin, (1968), Equilibrium behavior of population genetic models with non-random mating, Part II: Pedigrees, homozygosity, and stochastic models, J. App. Prob. 5: 487–566.

18. S. Karlin and F. M. Scudo, (1969), Assortative mating based on phenotype, II. Two autosomal alleles without dominance, Genetics 63: 499–510.

19. S. Karlin and M. W. Feldman, (1970), Linkage and selection: The two locus symmetric viability model, Theor. Pop. Biol. 1: 39–72.

20. S. Karlin, (1970), Lecture notes on mathematical genetics, Notes at Weizmann Institute.

21. S. Karlin and J. McGregor, (1971), On mutation selection balance for two-locus haploid and diploid populations, Theor. Pop. Biol. 2: 60–70.

22. S. Karlin and J. McGregor, (1972), Addendum to a paper of Ewen, Theor. Pop. Biol., March 1972.

23. S. Karlin, On some genetic models with an infinite number of types involving mutation selection balance, (to appear).

24. O. Kempthorne, An Introduction to Genetic Statistics, Wiley, New York, 1957.

25. J. F. C. Kingman, (1961), A mathematical problem in population genetics, Proc. Camb. Phil. Soc. 57: 574–582.

26. K. Kojima, (1970), Volume Edited "Mathematical Topics in Population Genetics". Biomathematics Vol. 1. Springer-Verlag, New York.

27. K. Kojima, and R. C. Lewontin, (1970), Evolutionary significance of linkage and epistasis, Mathematical Topics in Population Genetics, Vol. edited K. Kojima, Biomathematics Vol. 1. Springer-Verlag, New York.

28. R. C. Lewontin, (1968), The effect of differential viability on the population dynamics of *t* alleles in the house mouse, Evolution 22: 262–273.

29. F. M. Scudo, (1964), Sex population genetics, La Ricerca Scient. 34: 93–146.

30. F. M. Scudo and S. Karlin, (1969), Assortative mating based on phenotype I, Two alleles with dominance, Genetics 63: 479–498.

31. M. Slatkin, (1970), Selection and polygenic characters, Proc. Nat. Acad. Sci. 87–93.

32. C. Stern, (1960), An introduction to human population genetics, Freeman, San Francisco.

THE THEOREMS OF BONY AND BREZIS ON FLOW-INVARIANT SETS

R. M. REDHEFFER, University of California, Los Angeles

Throughout this note Ω is a domain in real Euclidean space E_n, $X(x)$ is a function on Ω to E_n, and F is a closed subset of Ω. We shall be concerned with trajectories of the vector field X, that is, with solutions of

$$\frac{dx}{dt} = X[x(t)], \quad x(t) \in \Omega.$$

The set F is *flow invariant* for X if every trajectory $x(t)$ which meets F at t_0 must remain in F for $t > t_0$. Thus, in the case of flow invariance,

$$x(t_0) \in F \Rightarrow x(t) \in F \text{ for } t_0 \leqq t < t_1,$$

where $[t_0, t_1)$ is the interval of existence for the trajectory through the point $x(t_0)$. When the solution does not exist beyond t_0, the condition is considered to be vacuously fulfilled.

Our objective is to generalize a remarkable theorem for flow-invariant sets that was recently obtained by Bony [2] and to show its relation to another theorem of Brezis [3]. The proofs here are simpler than those given hitherto, and the results are stronger. However, this paper is expository.

1. The theorems of Bony. Let $y \in F$ and let S be a sphere which has y on its boundary but does not contain any point of F in its interior. If S is centered at x, the vector $v(y) = x - y$ is normal to F at y in the sense of Bony. The following hypotheses involving v are used only at points y admitting a normal in this sense. In other words, if there is no sphere S as described above, the hypotheses are considered to be vacuously fulfilled.

For a given real-valued function δ, the upper left and right Dini derivates are respectively.

$$D^-\delta(t) = \limsup_{h\to 0+} \frac{\delta(t) - \delta(t-h)}{h}, \quad D^+\delta(t) = \limsup_{h\to 0+} \frac{\delta(t+h) - \delta(t)}{h}.$$

The lower Dini derivates D_- and D_+ are defined similarly, with lim inf instead of lim sup.

Professor Redheffer is known as a prolific mathematical analyst and an outstanding teacher. He received his Ph.D. at MIT under N. Levinson, held a Pierce Instructorship at Harvard, and has been at UCLA since. He spent an NSF year at Göttingen, a Fulbright year at Vienna and Hamburg, was visiting professor at both Hamburg and Berlin, and is guest professor at the Mathematisches Institut der Universität Karlsruhe. He is the author (with Sokolnikoff) of *Mathematics of Physics and Modern Engineering*. He was the mathematical consultant for the IBM-Eames exhibits "Mathematica" in Los Angeles, Seattle, and Chicago. He has worked extensively with William Johntz's Project SEED. *Editor.*

We say that a real-valued function ρ is a *uniqueness function* if the conditions

$$D^{-}\delta(t) \leqq \rho[\delta(t)], \quad D^{+}\delta(t) \leqq \rho[\delta(t)], \quad 0 < t < \varepsilon$$

together imply $\delta(t) = 0$, $0 < t < \varepsilon$, for every continuous function $\delta(t)$ satisfying

$$\delta(t) \geqq 0, \ \delta(0) = 0.$$

The uniqueness is required only for some positive ε.

THEOREM 1. (Bony). *Let X and F satisfy the following two conditions:*
(i) $(x-y)\cdot[X(x)-X(y)] \leqq |x-y|\rho(|x-y|)$ *for a uniqueness function* ρ;
(ii) $v(y)\cdot X(y) \leqq 0$ *whenever $v(y)$ is normal to F at y.*
Then F is flow-invariant for X.

Bony's theorem in its original form [2] is obtained when condition (i) is replaced by the familiar Lipschitz condition,

$$|X(x) - X(y)| \leqq K|x - y|, \ K \text{ constant}.$$

This corresponds to the choice $\rho(s) = Ks$, which is well known to be a uniqueness function in the above sense.

If Theorem 1 does not hold we can find t_0 such that $x(t_0) \in F$, but $x(t)$ is not in F on some interval $t_0 < t < t_1$ on which $x(t)$ exists. In all such cases we shall take $t_0 = 0$, as can be done without loss of generality. Let t be on $0 < t < t_1$ and let $\delta(t)$ denote the distance from $x(t)$ to F. Then

$$\delta(0) = 0, \ \ \delta(t) > 0 \ \text{ for } \ 0 < t < t_1.$$

For fixed t on $(0, t_1)$ let $x_h = x(t+h)$, let $x = x(t)$, and let $y \in F$ be a nearest point to x. Evidently

$$\delta(t+h) \leqq |x_h - y|, \ \delta(t) = |x-y|,$$

and hence, by the identity $a - b = (a^2 - b^2)/(a+b)$,

$$\delta(t+h) - \delta(t) \leqq \frac{|x_h - y|^2 - |x-y|^2}{|x_h - y| + |x - y|}. \tag{1}$$

The differential equation $dx/dt = X(x)$ gives

$$x_h = x + hX(x) + o(h).$$

If we compute $x_h - y$ from this and dot the result with itself, the numerator in (1) is found to be

$$2h(x-y)\cdot X(x) + o(h).$$

Dividing (1) by h and letting $h \to 0+$ therefore gives

$$D^{+}\delta(t) \leqq \frac{(x-y)\cdot X(x)}{|x-y|}. \tag{2}$$

The vector $v(y) = x - y$ is normal to F at y in the sense of Bony, and hence $(x-y) \cdot X(y) \leqq 0$. If this term is subtracted from the numerator in (2) the resulting inequality is

$$D^+\delta(t) \leqq \frac{(x-y) \cdot [X(x) - X(y)]}{|x-y|} \leqq \rho(|x-y|) = \rho[\delta(t)].$$

A more difficult argument, which we omit, gives a corresponding inequality for $D^-\delta(t)$. Since ρ is a uniqueness function, it follows that $\delta(t) = 0$, and this is a contradiction.

According to Bony the field X is *tangent to* F if $v(y) \cdot X(y) = 0$ for every $y \in F$ admitting a normal $v(y)$. In that case one can apply Theorem 1 as it stands and again with $-t$ replacing t. The result is the following, due also to Bony for the case $\rho(s) = Ks$:

THEOREM 2. (Bony). *Let X be tangent to F and let*

$$|X(x) - X(y)| \leqq \rho(|x-y|),$$

where ρ is a uniqueness function. Then any trajectory of $dx/dt = X(x)$ which meets F in one point must lie entirely in F.

The surprise in Theorems 1 and 2 is that F can fail to have a normal at a great many points, and it is by no means obvious *a priori* that the trajectory $x(t)$ could not escape from F at such a point. One of the main applications is to the sharp maximum principle [2], [5]. This application uses the full force of Bony's formulation, both as regards the one-sided condition (ii) and as regards the generality of the closed set F. At an opposite extreme, let F be the trace of a given solution-curve, $\bar{x}(t)$. The statement that $x(t) \in F$ is then the familiar uniqueness theorem for autonomous systems.

2. The theorem of Brezis. To state the next result let $|x, F|$ denote the distance from any point x to the closed set F. We then have:

THEOREM 3. *Let X and F satisfy the following two conditions*:
(i) $(x-y) \cdot [X(x) - X(y)] \leqq |x-y| \rho(|x-y|)$ *for a uniqueness function* ρ;
(ii) $\liminf_{h \to 0+} \dfrac{|y + hX(y), F|}{h} = 0$ *for each* $y \in F$.
Then F is flow invariant for X.

The condition (ii) is needed only at each y which possesses a normal in the sense of Bony. If there exists a trajectory satisfying

$$\frac{dx}{dt} = X(x), \quad x(0) = y,$$

then $x(h) = y + hX(y) + o(h)$ and the hypothesis (ii) is indistinguishable from

$$\liminf_{h\to 0+} \frac{|x(h),F|}{h} = 0.$$

This formulation bears an interesting relation to the conclusion, since the latter means that $|x(h),F| = 0$ for all $h \geqq 0$ on the interval of existence.

To prove Theorem 3, let v be normal to F in Bony's sense at $y \in F$ and let the sphere associated with v have center x, so that $v = x-y$. For $h \geqq 0$ it is convenient to set

$$\varepsilon(h) = |y + hX(y),F|. \tag{3}$$

Clearly

$$|x-y| \leqq |x,F| \leqq |x-y-hX(y)| + \varepsilon(h), \tag{4}$$

where the first inequality follows from the fact that the sphere associated with $v(y)$ is free of points of F, and the second follows from

$$|x,F| \leqq |x-\tilde{x}| + |\tilde{x},F| \tag{5}$$

with $\tilde{x} = y + hX(y)$. If the middle term is omitted from (4) and the resulting inequality is squared, we get

$$0 \leqq -2h(x-y)\cdot X(y) + o(h) + O[\varepsilon(h)].$$

Dividing by h and letting $h \to 0+$ through a suitable sequence, gives

$$X(y)\cdot(x-y) \leqq 0$$

which is Bony's condition (ii). Thus Theorem 3 follows from Theorem 1.

We want to formulate a weaker version of Theorem 3 which is very easy to prove, and yet generalizes the result of Brezis. To this end, ρ is called a *restricted uniqueness function* if the inequality

$$D_+\delta(t) \leqq \rho[\delta(t)], \quad 0 < t < \varepsilon,$$

implies $\delta(t) = 0$ for the same class of functions $\delta(t)$ as that considered above. Clearly, restricted uniqueness functions are also uniqueness functions.

THEOREM 4. (Brezis). *Let X and F satisfy the following two conditions*:
(i) $|X(x) - X(y)| \leqq \rho(|x-y|)$ *for a restricted uniqueness function* ρ;
(ii) $\liminf_{h\to 0+} \dfrac{|y+hX(y),F|}{h} = 0$ *for each* $y \in F$.
Then F is flow-invariant for X.

When $\rho(s) = Ks$ and when the lim inf in (ii) is replaced by lim, the result is Brezis' theorem in its original form [3]. Theorem 4 follows from Theorem 3, which is stronger both as regards the class $\{\rho\}$ and as regards the condition (i).

To deduce Theorem 4 from first principles, let $\delta(t)$ and $x(t)$ be as in the proof of Theorem 1, and define $\varepsilon(h)$ by (3). Then by (5)

$$\delta(t+h) \leqq \left| x(t+h) - y - hX(y) \right| + \varepsilon(h).$$

Since $x(t+h) = x + hX(x) + o(h)$ this gives

$$\delta(t+h) \leqq \left| \delta(t) + hX(x) - hX(y) \right| + o(h) + \varepsilon(h)$$

and hence

$$\delta(t+h) - \delta(t) \leqq h \left| X(x) - X(y) \right| + o(h) + \varepsilon(h).$$

Upon dividing by h and letting $h \to 0+$ through a suitable sequence, we get

$$D_{+}\delta(t) \leqq \rho[\delta(t)].$$

The conclusion follows at once.

Instead of considering the point $y + hX(y)$ as above, Brezis considers the point $x(h)$ on the trajectory satisfying

$$\frac{dx}{dt} = X(x), \qquad x(0) = y.$$

This seemingly minor alteration makes quite a difference, because the proof now depends on the existence of the trajectory through y and on its stability with respect to the initial value, y. (The first step of Brezis' proof invokes the stability inequality, which was not used here.) Existence and stability are available in the case $\rho(s) = Ks$ considered by Brezis, but are less immediate for general ρ.

3. Osgood functions. Discussion of the first-order equation for ρ involves knowledge of Dini derivates, and some of their properties are given now. In a general way, it can be said that these properties resemble those of ordinary derivatives. For instance, if f and $\phi \geqq 0$ are continuous then

$$D \int_0^{f(t)} \phi(s)ds = \phi[f(t)]Df(t), \tag{6}$$

where D stands for any one of the four derivates. The proof for D^- and D_- follows from

$$\frac{1}{h} \int_{f(t-h)}^{f(t)} \phi(s)ds = \frac{f(t)-f(t-h)}{h}\phi(\xi),$$

where ξ is between $f(t)$ and $f(t-h)$. This, in turn, is just the first mean-value theorem for integrals. Proof for D^+ and D_+ is similar.

As another illustration, suppose the continuous function g satisfies

$$Dg(t) < 1, \; 0 < t \leqq t_1; \; g(0) = 0, \tag{7}$$

where D is one of the derivates. Then $g(t) \leqq t$ on this interval. We give the proof for D_-; the case D_+ is a little harder. If the conclusion fails, the function $G(t) = g(t) - t$ attains a positive maximum at some point t, $0 < t \leqq t_1$. Thus $G(t-h) \leqq G(t)$ for each small positive h or equivalently,

$$\frac{g(t) - g(t-h)}{h} \geqq 1.$$

Hence the lim inf is also $\geqq 1$ and this is a contradiction.

A function $\rho(s)$ is an *Osgood function* if ρ is continuous, nonnegative, and if

$$\int_0^\eta \frac{ds}{\rho(s)} = \infty$$

for each small positive η. Since the meaning of the integral is not clear when 0 is a limit point of zeros of ρ, we agree that the above equation means

$$\lim_{\varepsilon \to 0+} \int_0^\eta \frac{ds}{\varepsilon + \rho(s)} = \infty. \tag{8}$$

In other words, the integral is interpreted in the sense of Lebesgue.

THEOREM 5. *Every Osgood function is a uniqueness function for each of the four Dini derivates, hence is usable for ρ in Theorems* 1–4.

The fact that Osgood functions are uniqueness functions is well known, but the following proof, based on [6] and [7], is simpler than proofs sometimes given. For $\varepsilon > 0$ define

$$g(t) = \int_0^{\delta(t)} \frac{ds}{\varepsilon + \rho(s)}.$$

If D denotes D_- or D_+, then by (6) and by $D\delta \leqq \rho(\delta)$,

$$Dg(t) = \frac{D\delta(t)}{\varepsilon + \rho[\delta(t)]} \leqq \frac{\rho[\delta(t)]}{\varepsilon + \rho[\delta(t)]} < 1.$$

Since $g(0) = 0$ we get $g(t) \leqq t$ by (7) and hence

$$\int_0^{\delta(t)} \frac{ds}{\varepsilon + \rho(s)} \leqq t_1, \quad 0 < t \leqq t_1. \tag{9}$$

If $\delta(t) = \eta > 0$ at some point t, this choice of t in (9) contradicts (8).

4. Further discussion of uniqueness. So far, we have required uniqueness for arbitrary continuous functions $\delta(t)$. However, the function $\delta(t)$ for which uniqueness is actually needed is somewhat restricted; it is the composition of a Lipschitzian function with the differentiable function $x(t)$. To see this, note that (5) as it stands

and (5) with x and $\tilde{x}$ interchanged gives

$$(10) \qquad |L(x) - L(\tilde{x})| \leqq |x - \tilde{x}|,$$

where $L(x) = |x, F|$. Since $\delta(t) = |x(t), F| = L[x(t)]$, the above remark is verified.

If X is locally bounded, then by (10)

$$|\delta(t) - \delta(\tilde{t})| \leqq |x(t) - x(\tilde{t})| \leqq M|t - \tilde{t}|,$$

where M is a bound for $|dx/dt| = |X(x)|$ in the relevant neighborhood, and hence, $\delta(t)$ is locally Lipschitzian. If, in addition, X is continuous, then $\delta(t) = o(t)$ as $t \to 0+$. This holds under Brezis' hypothesis whether X is continuous or not. To get it under Bony's hypothesis, note that the equation below (2) implies

$$(11) \qquad D^+\delta(t) \leqq |X(x) - X(y)|.$$

As $t \to 0+$ clearly $x \to x(0) \in F$, hence the nearest point y approaches $x(0)$ also, and the right side of (11) is less than ε near $0+$ for each positive ε. Applying (7) to $\delta(t)/\varepsilon$ gives $\delta(t) \leqq \varepsilon t$ near 0, as desired.

The reader familiar with uniqueness theorems of Kamke will know that the condition $\delta(t) = o(t)$ at $0+$ usually extends the class of functions ρ for which uniqueness holds. Accordingly, we call ρ a *generalized uniqueness function* if the conditions

$$(12) \qquad D^-\delta(t) \leqq \rho[\delta(t)], \quad D^+\delta(t) \leqq \rho[\delta(t)], \quad 0 < t < \varepsilon$$

imply $\delta(t) = 0$, $0 < t < \varepsilon$, for every function $\delta(t)$ on $0 \leqq t < \varepsilon$ which satisfies

$$\delta(t) \geqq 0, \ \delta(t) \in \mathrm{Lip}\,1, \quad \lim_{t \to 0+} \frac{\delta(t)}{t} = 0.$$

So far we have required that $dx/dt = X(x)$ hold for all t. It is usually sufficient, however, to have $x(t)$ continuous and to have the differential equation hold except perhaps on a countable set. When such is the case it is said that the differential equation holds $\mathrm{mod}\, E$.

By considering the integral of a Cantor function one sees that the hypothesis $\mathrm{mod}\, E$ cannot be replaced by a similar hypothesis $\mathrm{mod}\, N$, where N denotes an arbitrary null set. However, the extension can be made if x is required to be absolutely continuous. In that case the differential equation can be interpreted as an integral equation,

$$x(t) = \int_{t_0}^{t} X[x(s)]ds + x(t_0).$$

Clearly $\delta(t)$ is continuous if $x(t)$ is. To check for absolute continuity one would consider

$$|x(t_1) - x(t_2)| + |x(t_3) - x(t_4)| + \cdots + |x(t_{m-1}) - x(t_m)| \leqq \eta.$$

This gives a similar inequality for $\delta(t) = L[x(t)]$ and hence, L maps the absolutely continuous functions on E_n into absolutely continuous functions on E_1. It is also true that the above analysis gives (12) at each point t, where $dx/dt = X(x)$. Hence if the latter holds $\bmod E$ or $\bmod N$, as the case may be, so does the former.

It is left for the reader to formulate what is meant by a uniqueness function $\bmod E$ or $\bmod N$. The results of this discussion are then summarized as follows:

THEOREM 6. *In Theorems* 1–3 *suppose the hypothesis is changed in one of the following three ways*:

(i) X is *continuous and* ρ *is a generalized uniqueness function*; *or*

(ii) $dx/dt = X(x) \bmod E$, *and* ρ *is a uniqueness function* $\bmod E$; *or*

(iii) $dx/dt = X(x) \bmod N$, *and* ρ *is a uniqueness function* $\bmod N$.

Then the conclusions still hold.

The most important special case is given by the following:

THEOREM 7. *The conclusions of Theorems* 1–3 *hold for every Osgood function* ρ, *even if the differential equation* $dx/dt = X(x)$ *is given only* $\bmod E$ *or* $\bmod N$.

5. Acknowledgment. The content of Sections 2–4 was enriched by several conversations with Professor M. Crandall, and this influence is gratefully acknowledged.

The preparation of this paper was supported in part by NSF Grant No. GP–13377.

References

1. George Aumann, Reelle Funktionen, Springer-Verlag, (1969) 213–228.

2. Jean-Michel Bony, Principe du Maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés, Ann. Inst. Fourier, Grenoble, 19 (1969) 277–304.

3. Haim Brezis, On a characterization of flow-invariant sets, Comm. Pure App. Math., 223 (1970) 261–263.

4. T. H. Hildebrandt, Introduction to the theory of integration, Academic Press, New York, (1963) 347–357.

5. C. Denson Hill, A sharp maximum principle for degenerate elliptic-parabolic equations, Indiana University Math. J., 20 (1970) 213–229.

6. R. M. Redheffer, Differential and integral inequalities, Proc. Amer. Math. Soc., (1964) 715–716.

7. ———, Bemerkungen über Differentialungleichungen bei abzählbaren Ausnahmemengen, Numerische Math., 9 (1967) 437–445.

8. Wolfgang Walter, Differential and integral inequalities, Springer-Verlag, 1970, Chapter II.

WHAT IS A REAL NUMBER?

JOHN MYHILL, State University of New York at Buffalo

In this paper I shall try by examples to give some of the feeling of constructive mathematics. I shall adopt the point of view of Bishop, which is in many ways clearer than that of Brouwer, the originator of constructive mathematics. I shall consider the notion of real numbers from a constructive point of view. This point of view requires that any real number can be *calculated*. It does not believe in the existence of any object which has not been constructed. We shall explain various senses in which it can be said that a real number has been constructed, and explain why some of these are unsuitable for the purpose of developing analysis constructively.

As a first approximation, let us say that a real number has been constructed if a rule has been given which enables us to compute its nth decimal place for any positive integer n. The notion of a "rule" is a primitive one in constructive mathematics, but it must be understood that the application of a rule is a mechanical matter; no intelligence is involved. In particular we may think of a digital computer, which given any positive integer n, will print out the number $f(n)$, as defining the rule f. In fact nobody has ever given an example of a function from positive integers to positive integers which can be calculated in a mechanical way, other than those which can be calculated by suitably idealized digital computers—the so-called *recursive functions*. Thus in practice it might suffice to identify rule-like functions of natural numbers with recursive functions. This identification, however, does not in our opinion belong to mathematics but to philosophy, and we shall abstain from making it. We therefore take the notion of a rule as an undefined one; in practice we seem to be able always to recognize when a mechanical process has been described.

From the constructive point of view, the only functions which exist are those which have been constructed; that is, functions for whose evaluation a rule has been given. For example, if we define

$$f(x) = \begin{cases} 0, \text{ if } a^x + b^x \neq c^x \text{ for all} \\ \text{integers } a, b, c > 0, \\ 1, \text{ if } a^x + b^x = c^x \text{ for some} \\ a, b, c > 0, \end{cases}$$

John Myhill received his Harvard Ph.D. under W. V. Quine. He has since held positions in philosophy and mathematics departments at Vassar, Temple, Yale, Chicago, Berkely, Stanford, Illinois, and presently SUNY at Buffalo. He held a Guggenheim Fellowship while at Chicago, and he spent three fellowship years at the Institute for Advanced Study. He spent the last academic year as a visitor at the Univ. of Leeds, England. He is the author of numerous papers on recursive functions, foundations of mathematics, and computer science; also of *Recursive Equivalence Types* (with J. C. E. Dekker, Univ. of Calif. Press, 1960) and of *Intuitionism and Proof Theory* (co-editor with Kino and Vesley, North Holland, 1969). *Editor.*

we have defined, from the classical point of view, a function. However, from a constructive point of view this does not constitute a definition of a function, because no directions have been given for computing it.

Because of the restriction to rule-like functions, we shall henceforth use the words 'rule' and 'function' interchangeably. Thus we shall not regard f just defined as being a function at all.

Our first attempt at explaining what is meant by a real number is then as follows: α is a real number if a rule has been given to compute the nth decimal place of α. Thus a real number α can be identified with a function ϕ from non-negative integers to integers, where $\phi(0)$ is the integer part of α and where for $n > 0$, $\phi(n) \in \{0, \cdots, 9\}$. We shall denote the set of real numbers in the sense of this definition by R_d (d for "decimal").

Although most of the real numbers encountered in analysis (for example all the algebraic numbers, and the transcendental numbers e and π) are constructible in this sense, we shall show that the set R_d is not suitable as a foundation for analysis. In fact we prove the following disagreeable thing:

THEOREM 1. *The set R_d of real numbers possessing a decimal expansion is not closed under addition, i.e., there are numbers α, $\beta \in R_d$ such that the number $\alpha + \beta$ is not in R_d.*

Before I prove this, I must explain the constructive sense of the word "not". This is used in *historical* sense; that is, to say that a proposition is not true means that no one has yet proved it. From the constructive point of view, just as nothing exists until it has been constructed, so no proposition is true until it has been proved. Constructivists reject the idea that in some platonic realm a T or an F has been placed beside each mathematical proposition P, independently of whether anyone knows whether P is true. There is another constructive notion resembling "not", called *absurdity*: P is called *absurd* if the assumption of P yields a contradiction. The notion of absurdity shares some of the properties of the classical "not", but it does not, for example, satisfy the law of excluded middle; it is simply untrue that for every proposition P, P has either been proved or shown to be absurd (contradictory). The law of excluded middle, "P or not P" in the classical sense, appears to the constructivist to be a piece of mythology; it says that in some non-material world, truth-values have already been assigned to all propositions, independent of human mathematical activity. Constructivists cannot make sense of this third kind of "not"; a truth that nobody knows how to prove makes as little sense to the constructivist as a real number that nobody knows how to calculate.

In Theorem 1, "not" is used in the historical sense. We shall give two numbers α and β such that each of them can be computed to any required number of decimal places, while yet nobody knows even the first decimal place of $\alpha + \beta$. To prove this (historical) assertion, we shall use our (historical) ignorance of the behavior of the

decimal expansion of π. Specifically, nobody knows whether a sequence 5555 occurs in that expansion. If such a sequence occurs beginning at the kth place, and if it is the first such sequence, then k is called the *critical number* (of π). Nobody knows whether such a number exists, and nobody knows whether (if it exists) it is even or odd. Further, given any nonnegative integer n, one can evidently determine whether n is critical or not; all one has to do is compute the first $n + 3$ decimal places of π.

Now I give directions for computing the decimal expansions of the numbers α and β.

To compute α we write down

$$.33333\dots\dots\dots$$

and continue writing 3 unless we reach some *odd* place, $2n + 1$, such that $2n + 1$ is the critical number of π. In that case we write a 4 at the $2n + 1$-st place and ever afterwards.

Thus if the critical number k of π is odd, $\alpha > \frac{1}{3}$, but if k is even or does not exist, $\alpha = \frac{1}{3}$.

To compute β we write down

$$.66666\dots\dots\dots$$

and continue writing 6 unless we reach some *even* place, $2n$, such that $2n$ is the critical number of π. In that case we write a 5 at the $2n$th place and ever afterwards.

Thus if the critical number k of π is even, $\beta < \frac{2}{3}$, but if k is odd or does not exist, $\beta = \frac{2}{3}$; ("k does not exist" means "no 5555 occurs in the decimal expansion of π").

We have

if k is even $\alpha = \frac{1}{3}$, $\beta < \frac{2}{3}$, $\alpha + \beta < 1$;

if k is odd $\alpha > \frac{1}{3}$, $\beta = \frac{2}{3}$, $\alpha + \beta > 1$; and

if k does not exist, $\alpha = \frac{1}{3}$, $\beta = \frac{2}{3}$, $\alpha + \beta = 1$.

Now suppose we could write down even one place of the decimal expansion of $\alpha + \beta$. Then

$$\text{if } \alpha + \beta \text{ begins } 1 \cdot \cdots, \text{ then } \alpha + \beta \geqq 1$$

and if k exists, it is odd; while

$$\text{if } \alpha + \beta \text{ begins } \cdot 9 \cdots, \text{ then } \alpha + \beta \leqq 1$$

and if k exists, it is even.

Thus if we could compute even one place of $\alpha + \beta$, we could prove one of the two propositions "if k exists, it is odd" or "if k exists, it is even." That is, we could either prove "if 5555 occurs in π, its first occurence begins at an odd place," or else we could prove "if 5555 occurs in π, its first occurence begins at an even place." But we have not proved either of these two propositions; thus we cannot write down

even one decimal place of $\alpha + \beta$, even though we can write down all the places of α and β. This completes the proof of Theorem 1.

Now we consider another possible approach to real numbers. One can object to the above proof that it is artificial; it uses the numbers α and β that are not *located* with respect to the rationals. To say a real number λ is located with respect to the rationals is to say that we can decide, for every rational number r, which of the three alternatives $\lambda < r$, $\lambda = r$, $\lambda > r$, holds. Thus α is not located with respect to $\frac{1}{3}$, β is not located with respect to $\frac{2}{3}$, and $\alpha + \beta$ is not located with respect to 1. We shall also require that we know an integer upper bound M on $|\lambda|$. This enables us to compute the decimal expansion of any located real number λ. For we first compare λ with each of the integers

$$-M,\ -M+1,\cdots,0,\cdots,M-1,M$$

to get the whole number part of λ, say q; then we compare λ with each of $q+\frac{1}{10}$, $q+\frac{2}{10},\cdots,q+\frac{9}{10}$, to get the first place after the decimal point, and so on. The situation is as follows:

THEOREM 2. *Let R_l (l for "located") denote the set of all located real numbers. Then $R_l \subset R_d$, but the converse does not hold.*

$R_l \subset R_d$ we have just proved. To disprove $R_d \subset R_l$, we must give a number with a decimal expansion which is not located with respect to the rationals. The number α of the preceding theorem is such a number. For we showed how to compute its successive decimal places, but we have not proved any of the three propositions "$\alpha < \frac{1}{3}$," "$\alpha = \frac{1}{3}$," or "$\alpha > \frac{1}{3}$." ($\alpha < \frac{1}{3}$ is absurd since every digit of α is either 3 or 4; $\alpha = \frac{1}{3}$ would imply "if k exists it is even," and $\alpha > \frac{1}{3}$ would imply "k exists and is odd." But we have not proved either of these propositions.) Hence $\alpha \in R_d - R_l$.

The condition of being located is therefore strictly stronger than that of having a decimal expansion. Furthermore, most of the real numbers encountered in analysis are located—the algebraic numbers for example, and the numbers e and π, as was shown by Goodstein. By way of illustration the number $\sqrt{2}$ is located; for to determine whether a rational r is $<$ or $>$ $\sqrt{2}$ ($=$ of course is impossible), we simply ask first if $r \leqq 0$; if it is, then $r < \sqrt{2}$, if not we compute r^2 and ask if $r^2 <$ or > 2. In fact we can prove a stronger property of $\sqrt{2}$ which we shall need in the sequel.

THEOREM 3. *For any rational number r, we can compute a number n_r such that* $|r - \sqrt{2}| > 1/10^{n_r}$. (This means that the decimal expansion of r differs from that of $\sqrt{2}$ at or before the n_rth place.)

Proof. We have

$$|r - \sqrt{2}| \geqq \big||r| - \sqrt{2}\big| = \frac{|r^2 - 2|}{|r| + \sqrt{2}} > \frac{|r^2 - 2|}{|r| + 2}.$$

So pick n_r so large that $1/10^{n_r} < |r^2 - 2|/(|r| + 2)$.

It will probably be felt that any reasonable number is located, and that the fact that R_d is not closed under addition results from the fact that numbers like α and β in Theorem 1 are not located. This might incline us to define computable real numbers λ as located rather than decimally expandible real numbers; formally, as pairs (N,f), where $N > |\lambda|$ and where for each rational $r, f(r) = 0$, 1, or 2, according as $\lambda <, =,$ or $> r$. But this too will not do since we have more trouble:

THEOREM 4. *R_l is not closed under addition, i.e., there exist numbers γ, $\delta \in R_l$ such that $\gamma + \delta \notin R_l$.*

Proof. Let $\gamma \equiv \sqrt{2}$. We do not know whether 5555 occurs in the decimal expansion of $\sqrt{2}$. Define 'critical number of $\sqrt{2}$' as we defined 'critical number of π' before. To compute δ, write down the decimal expansion of $\sqrt{2}$, except that if n is the critical number of $\sqrt{2}$, we write 0 at the nth place and thereafter. Clearly $\gamma \in R_l$. Clearly also $\gamma + \delta \notin R_l$. For if $\gamma + \delta = 0$, $\delta = \sqrt{2}$ and no 5555 occurs in the decimal expansion of $\sqrt{2}$; while if $\gamma + \delta < 0$, such a 5555 does occur. Thus if $\gamma + \delta$ were located with respect to 0 we could determine whether $\sqrt{2}$ possesses a critical number, which we cannot. It remains to prove $\delta \in R_l$.

Let then a rational number r be given; we must show how to decide $r < \delta$, $r = \delta$, or $r > \delta$. First find if $r > \sqrt{2}$ or $r < \sqrt{2}$.

CASE I. $r > \sqrt{2}$. Then certainly $r > \delta$, for $\delta \leqq \sqrt{2}$.

CASE II. $r < \sqrt{2}$. By Theorem 3 we can find n such that r and $\sqrt{2}$ differ at or before the nth decimal place. Let n_0 be the least such n. The n_0th place of r is less than n_0th place of $\sqrt{2}$. Then if (SUBCASE II.1) there is no critical number of $\sqrt{2} \leqq n_0$, $\sqrt{2}$ and δ agree for their first n_0 places and $r < \delta$. If on the other hand (SUBCASE II.2) there is a critical number $\leqq n_0$, we can compute β exactly and compare it with r directly. In any case we can decide whether $r <, =,$ or $> \delta$ and so $\delta \in R_l$.

So we cannot use R_l as a foundation for analysis. We now give another definition which avoids the above difficulties. A *finite decimal* is a number of the form $a/10^b$, where a is an integer and b is a nonnegative integer; a real number ρ is called *decimally approximable* ($\rho \in R_{da}$) if given any rational $\varepsilon > 0$ we can find a finite decimal d with $|\rho - d| < \varepsilon$. This is wider than either of the preceding notions.

THEOREM 5. *$R_d \subset R_{da}$, but the converse implication does not hold.*

Proof. Let $\rho \in R_d$. Then by definition we can compute any desired number of places of the decimal expansion of ρ. To approximate it within $1/10^n$ we need only compute $n + 1$ places; hence $\rho \in R_{da}$. To refute the converse observe that the sum of two elements of R_{da} is again in R_{da}. For if d_1 and d_2 are decimal $\varepsilon/2$-approximations to ρ_1 and $\rho_2 \in R_{da}$, then

$$|(d_1 + d_2) - (\rho_1 + \rho_2)| \leqq |d_1 - \rho_1| + |d_2 - \rho_2| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so that $d_1 + d_2$ is a decimal ε-approximation to $\rho_1 + \rho_2$. Hence $\rho_1 + \rho_2 \in R_{da}$. Now

let α, β be as in Theorem 1, α, $\beta \in R_d$ but $\alpha + \beta \notin R_d$. Then α, $\beta \in R_{da}$ and so $\alpha + \beta \in R_{da} - R_d$.

Thus the decimally approximable real numbers form a more likely candidate as a foundation for constructive analysis than either R_d or R_l. The following theorem confirms this impression:

THEOREM 6. *R_{da} is a field.*

We have just proved that R_{da} is closed under addition; as an example of the verification of the remaining field postulates we shall prove that it is closed under multiplication. Let ρ_1, $\rho_2 \in R_{da}$: we seek a decimal ε-approximation to $\rho_1\rho_2$. We first compute (from 1-approximations to ρ_1 and ρ_2) a number $M > \max(|\rho_1|, |\rho_2|)$. Now find $\varepsilon/2M$-approximations d_1, d_2 to ρ_1, ρ_2 respectively, with $|d_1|, |d_2| < M$. We have

$$\begin{aligned} |d_1 - \rho_1|, |d_2 - \rho_2| &< \varepsilon/2M \\ |d_1d_2 - \rho_1\rho_2| &= |d_1(d_2 - \rho_2) + \rho_2(d_1 - \rho_1)| \\ &< M|d_2 - \rho_2| + M|d_1 - \rho_1| \\ &< M(\varepsilon/2M) + M(\varepsilon/2M) = \varepsilon, \end{aligned}$$

so that d_1d_2 is an ε-approximation to $\rho_1\rho_2$.

In verifying the field postulates, we have to make sure that the statement of some of them makes constructive sense. For example, in the postulate

$$(*) \qquad x \neq 0 \rightarrow (\exists y)(xy = 1)$$

we must be careful to give the right meaning to the hypothesis $x \neq 0$. It is easy to construct a number which is neither $<, =$, or > 0. For example, the number $\gamma + \delta$ in Theorem 4 is such a number (recall that if $\gamma + \delta = 0$, no 5555 occurs in the decimal expansion of $\sqrt{2}$; if $\gamma + \delta < 0$, such a 5555 does occur, while $\gamma + \delta > 0$ is absurd). Now $x \equiv \gamma + \delta \in R_{da}$, but x is neither $<, =$, or > 0. How are we to construe $(*)$ for such an x? The correct version is: If x is *separated* from zero, i.e., if a rational number r with $0 < r < |x|$ is known, then x possesses a reciprocal. This notion of separation is an example of how constructive mathematics (except in counter-examples) normally replaces negative statements by positive ones.

It may come as a surprise to some to learn that R_{da} is a *complete* field, in the sense that if $\{\rho_i\}$ is a sequence of elements of R_{da} such that for every $\varepsilon > 0$ we can compute N_ε with

$$|\rho_i - \rho_j| < \varepsilon \qquad (i,j > N_\varepsilon),$$

then we can construct a number $\lim \rho \in R_{da}$ satisfying

$$(\forall\varepsilon)(\exists M_\varepsilon)(\forall i > M_\varepsilon)|\rho_i - \lim \rho| < \varepsilon.$$

The proof is in fact a rather straightforward computation with ε's and δ's.

Of course R_{da} is not an *ordered* field; we just saw an example of an element $x \equiv \gamma + \delta$ of R_{da} which was neither $>, <$, or $= 0$. However, R_{da} is closed with

respect to all the usual functions connected in analysis, and indeed is sufficiently like the classical continuum that Bishop has made it the foundation of his book on constructive analysis. What is more remarkable is that the arguments of his book (not the counterexamples, but the theorems) are to an unexpected extent scarcely different from the classical ones. When they differ, they surpass the classical ones in precision and numerical content; for example, the proofs of existence always contain a method for approximating the number asserted to exist.

I conclude with two remarks of a more specialized nature. Firstly, I would like to make precise the difference between *constructive* analysis and *recursive* analysis. What we have been doing is constructive analysis; it admits no real numbers other than computable ones and no methods of proof other than constructive ones, and the notion of "computable function" or "rule" is a primitive one. Recursive analysis (e.g., in the sense of Klaua) on the other hand, is the study, by whatever means one wishes, of a certain classically defined subset of the real numbers, called the recursive reals. "Computable" is simply a synonym for "recursive" and is a defined idea. From the point of view of what I call recursive analysis, the sets R_d, R_l and R_{da} are all the same, but the proof that they are the same is non-constructive.

My last remark concerns the *formalization* of the remarks in this paper. If one takes a two-sorted theory, with variables for natural numbers and computable functions, and postulates, for the former, Peano's axioms and (primitive) recursive definition and for the latter, simply the axiom of choice

$$(\forall x)(\exists y)A(x,y) \rightarrow (\exists f)(\forall x)A(x,f(x));$$

and if the underlying logic is taken to be the intuitionistic predicate calculus, I think one has an adequate foundation for the constructive theory of real numbers. (Of course that is not the whole of constructive analysis; for the theory of functions of a real or complex variable one needs functionals of higher types for which one also postulates axioms of choice and the possibility of primitive recursive definition. But for our purposes it is enough to consider just the simple two-sorted theory mentioned.) Note that the notion of "recursive" or "computable" function does not appear at all; the function-variables range *only* over computable functions.

How, finally, is one to formalize in this theory the counter-examples we have been discussing? One possibility is to adjoin *rules of rejection* as well as rules of proof; for example let $P(x)$ denote 'x is the critical number of π', then we postulate $Px \wedge Py \rightarrow x = y$,

$$P(x) \vee \neg P(x)$$

(P is decidable) and assert that both of the formulas $(\forall x)(Px \rightarrow x \text{ is even})$ and $(\forall x)(Px \rightarrow x \text{ is odd})$ are to be rejected (as not yet proved). The rules of rejection are: if $A \vdash B$ and B is rejected, then A is rejected: if A and B are rejected, so is $A \vee B$; if $A(x)$ is rejected (for all x) so is $(\exists x)A(x)$. On this basis we can formally prove that the two inclusions $R_d \subset R_l$ and $R_{da} \subset R_d$ are rejected.

ADDENDUM TO "EMMY NOETHER"

C. H. KIMBERLING, University of Evansville

Professor Freeman J. Dyson of The Institute for Advanced Study has written me concerning the statement in "Emmy Noether" (this MONTHLY, 79 (1972) 136–149) that the letter written by Einstein to *The New York Times* was "inspired, if not written, by Dr. Hermann Weyl." Professor Dyson discussed this statement with Miss Dukas, Einstein's former secretary, who is presently in charge of the Einstein archive at The Institute for Advanced Study.

I quote from Professor Dyson's letter:

Miss Dukas has the original German draft of the letter. She confirms that this was written by Einstein himself at the request of Weyl. She does not remember whether Weyl or somebody else afterwards translated it into English.

Miss Dukas also has a letter from Einstein to Hilbert dated May 24, 1918, including the following passage:

"Gestern erhielt ich von Frl. Noether eine sehr interessante Arbeit ueber Invariantenbildung. Es imponiert mir, dass man diese Dinge von so allgemeinem Standpunkt uebersehen kann. Es haette den Goettinger Feldgrauen nichts geschadet, wenn sie zu Frl. Noether in die Schule geschickt worden waeren. Sie scheint ihr Handwerk zu verstehen!"

Here "Feldgrauen" is slang for "Warriors." From the letter you can see that, while it may be true that Einstein and Emmy Noether never met (Miss Dukas is not sure about this), Einstein certainly knew her work well and understood its importance early and at first hand.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

ON THE DIFFEOMORPHISMS OF EUCLIDEAN SPACE

W. B. GORDON, Mathematics Research Center, Naval Research Laboratory

1. A differentiable map f between manifolds is said to be a diffeomorphism if the map is one-one and onto and its inverse is also differentiable. A continuous map f is said to be proper if $f^{-1}(K)$ is compact whenever K is compact. Hence, to say that a continuous map f from R^N to R^N is proper means that $|x| \to \infty$ implies $|f(x)| \to \infty$.

THEOREM A. *A C^1 map f from R^N to R^N is a diffeomorphism if and only if f is proper and the Jacobian* $\det(\partial f_i/\partial x_j)$ *never vanishes.*

This theorem goes back at least to Hadamard [2, 3, 4], but it does not appear to be "well-known". Indeed, I have found that most people do not believe it when they see it and that the skepticism of some persists until they see two proofs. Now according to the Implicit Function Theorem, the non-vanishing of the Jacobian implies that f is a local homeomorphism, (i.e., that each x in R^N has an open neighborhood which is mapped homeomorphically by f onto an *open* subset of R^N), but this condition by itself does not insure that f is either one-one or onto. (Standard example: the map $(x_1, x_2) \to (e^{x_1}\cos x_2, e^{x_1}\sin x_2)$.) What seems to be surprising is that the addition of the hypothesis that f is proper guarantees that f is both one-one *and* onto.

A generalization of Theorem A to manifolds is provided by the following theorem, also known to Hadamard:

THEOREM B. *Let M_1 and M_2 be connected, oriented N-dimensional manifolds of class C^1, without boundary, and suppose that M_2 is simply connected. Then a C^1 map f from M_1 to M_2 is a diffeomorphism if and only if f is proper and the Jacobian of f never vanishes.*

REMARK. The simple connectivity of M_2 is necessary to insure that f is one-one, (e.g., the map $\exp(i\theta) \to \exp(2i\theta)$ which wraps the unit circle around itself twice, is proper and has non-vanishing Jacobian). But if this condition is removed, one can still conclude that f is onto. This is well known; see Section 3 below for references.

We shall give two independent proofs of Theorems A and B, which we hope will be agreeable to modern tastes. The proof of Theorem A will be confined to the case $f \in C^2$, but involves nothing more than the elementary stability theory of differential equations. The proof of Theorem B, which of course provides a proof of Theorem A in the C^1 case, is short but involves the use of some topological ideas, viz., the notion of a *universal covering space* of a topological space and the *topological degree* of a map.

REMARK. The "only if" parts of the theorems are easy to prove. For let f be a diffeomorphism. Then the inverse of f, being continuous, must map compact sets into compact sets. This shows that f is proper. Moreover, from the multiplicative property of Jacobians, it follows that the product of the Jacobian of f with that of its inverse (evaluated at the appropriate points) is unity. Hence the value of the former can never be zero.

2. Proof of Theorem A for the C^2 Case. We have to show that f is one-one and onto, i.e., that f has an inverse. Once this is accomplished, the Implicit Function Theorem will guarantee that the inverse is of class C^1.

(a) *The map f is onto.* Obviously, it suffices to show that $f(x) = 0$ for at least one x in R^N. Let

$$F(x) = \tfrac{1}{2}|f(x)|^2 = \tfrac{1}{2}\,\Sigma(f_r(x))^2,$$

so that

$$\frac{\partial F}{\partial x_i} = \Sigma f_r \frac{\partial f_r}{\partial x_i}.$$

Hence, from the non-vanishing of the Jacobian, $\nabla F(x) = 0$ if and only if $f(x) = 0$ if and only if $F(x) = 0$.

We now proceed to locate the zeros of F by the method of steepest descent. Consider the differential equation

$$(*) \qquad \frac{dx}{dt} = -\nabla F(x(t)).$$

Let $x = x(t)$ be a solution to $(*)$ with arbitrary initial condition. Along a solution curve, $F(x(t))$ is non-increasing since

$$\frac{dF}{dt} = \nabla F \cdot \frac{dx}{dt} = -|\nabla F|^2.$$

But F is proper since f is proper, so that the solution $x = x(t)$ remains in some compact set as t varies over any interval $[0, \omega)$ for which a solution is defined. This implies that solutions are defined for all $t \geqq 0$. Moreover, dF/dt cannot be bounded away from zero, since otherwise $F(x(t))$ would eventually become negative. Therefore, $\nabla F(x(t_n)) \to 0$ for some sequence $\{x(t_n)\}$. Again using the propriety of F, one can extract a convergent subsequence $x(t_{n'})$; $x(t_{n'}) \to p$, and $\nabla F(p) = 0$ by the continuity of ∇F. That is, we have obtained a solution p to $f(p) = 0$.

REMARK 1. This argument is standard. What we have done is to show that F satisfies "Condition C" of Palais and Smale. Cf. [7].

REMARK 2. We could have given a simpler proof for this part of the theorem, but have used the above argument for reasons that will become obvious later. The alternate proof follows: Let c be a number greater than some value of F. Then $F^{-1}[0, c]$ is non-empty, and compact since F is proper. Therefore, F attains a minimum at some point p in $F^{-1}[0, c]$, and $F(p)$ is the smallest value of F on the entire space R^N. Hence $\nabla F(p) = 0$, and therefore $f(p) = 0$.

The remainder of this section is devoted to a proof that f is one-one, i.e., there is only one solution to $f(x) = 0$. Let $S = f^{-1}(0)$.

(b) *S has only a finite number of elements.* S is compact since f is proper. Hence if S contained an infinite number of elements, there would exist at least one accumulation point q. But the non-vanishing of the Jacobian of f at q implies that f is one-one in a neighborhood of q.

Let $S = \{p_1, \cdots, p_n\}$

(c) *Each p_i in S has a neighborhood U_i such that any solution $x = x(t)$ to $(*)$ which enters U_i remains in U_i, and in fact converges to p_i as $t \to +\infty$. I.e., each p_i*

is an asymptotically stable critical point of the system $(*)$. For F is a Lyapunov function for the system $(*)$ at each p_i. Specifically, along a solution curve $x = x(t)$ we have $dF/dt \leqq 0$, and equality holds only if $x = x(t)$ is a trivial solution $x(t) \equiv p_i$.

Let W_i be the set of all q in R^N such that the solution $x = x(t)$ to $(*)$ with initial condition $x(0) = q$ satisfies $x(t) \to p_i$ as $t \to +\infty$.

(d) $R^N = \cup W_i$. This has already been proved in Parts (a) and (c).

(e) *Each of the* W_i *is open*. This is a consequence of the continuity of solutions with respect to initial conditions. Choose $\varepsilon > 0$ such that each ball with radius 2ε centered at p_i is contained in U_i. Suppose $q \in W_i$, and let $x = x(t)$ be the solution with $x(0) = q$. Then $|x(T) - p_i| < \varepsilon$ for some $T > 0$. Let $y = y(t)$ be the solution with $y(0) = q'$. Then by making $|q - q'|$ sufficiently small we can insure that $|x(T) - y(T)| < \varepsilon$, so that $y(T)$ is in U_i. But from Part (c), this implies that q' lies in W_i.

Now, putting everything together, we see that R^N is the union of a finite number of mutually disjoint, non-empty open subsets W_i. Hence, there exists only one. I.e., we have shown that $f^{-1}(0)$ is a single point.

3. Proof of Theorem B. The fact that f is onto is well-known and easy to prove once the basic properties of the topological degree of maps have been established. (See [1], [6], [8], [9].)

We shall prove now that f is one-one. The manifold M_2, being simply connected, is its own universal covering space. Hence, it suffices to show that f is a covering, i.e., we have to show that every q in M_2 has an open neighborhood V such that $f^{-1}(V)$ consists of disjoint open sets mapped homeomorphically by f onto V. This is accomplished by modifying a construction given by Milnor in [5, p. 8].

Since f is a proper local homeomorphism, $f^{-1}(q)$ consists of only a finite number of points, say, $p_1, \cdots, p_n$. Let K be a compact neighborhood of q. Then $f^{-1}(K)$ contains disjoint open neighborhoods U_i of the points p_i, which are mapped homeomorphically by f onto open neighborhoods of q. Let

$$V = (f(U_1) \cap \cdots \cap f(U_n)) - f[f^{-1}(K) - (U_1 \cup \cdots \cup U_n)].$$

We leave it to the reader to verify that V satisfies the desired conditions, and this concludes the proof.

My thanks are due to Professor Melvyn S. Berger for bringing my attention to the works of Hadamard cited in the references.

References

1. M. S. Berger and M. S. Berger, Perspectives in Non-Linearity, Benjamin, New York, 1968.
2. J. Hadamard, Sur les transformations planes, C. R. Acad. Sci. Paris, 142 (1906), 74.
3. ———, Sur les transformations ponctuelles, Bull. Soc. Math. France, 34 (1906) 71–84; Oeuvres, pp. 349–363.
4. ———, Sur les correspondances ponctuelles, Oeuvres, pp. 383–384.

5. J. Milnor, Topology from the Differentiable Viewpoint, U. Press of Virginia, Charlottesville, Va., 1965.

6. A. Nijenhuis and R.W. Richardson, Jr., A theorem on maps with non-negative Jacobians, Mich. Math. J., 9 (1962) 173–176.

7. R. S. Palais, Morse theory on Hilbert manifolds, Topology, 2 (1963) 299–340.

8. S. Sternberg, Lectures on Differential Geometry, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

9. S. Sternberg and R. G. Swan, On maps with nonnegative Jacobian, Mich. Math. J., 6 (1959) 339–342.

ON THE UNION OF CLOSED SETS OF A FINITE DIMENSIONAL VECTOR SPACE

D. E. RADFORD, Lawrence University

1. Introduction. Let k be an infinite field. In this note we discuss the union of closed sets (for example, subspaces) of a finite-dimensional vector space over k. We show that a finite-dimensional vector space over k is not the union of a family of proper closed subsets provided the cardinality of the family is not too large. As one application we find a general condition under which polynomial functions (in particular, functionals) are distinguished by a single vector.

2. The Zariski topology. For any vector space V over a field k, let $M(V,k)$ be the algebra of functions $f: V \to k$ under point-wise multiplication, and denote by $A(V)$ the subalgebra generated by the linear functionals V^*. If $f \in A(V)$, let $V(f) = \{v \in V: f(v) = 0\}$ be the zero-set of f, and for any non-void subset I of $A(V)$, let $V(I) = \cap \{V(f): f \in I\}$. The sets $V(I)$ are the closed sets of the *Zariski topology* on V. Let $T: V \to W$ be linear. Then $T^*: A(W) \to A(V)$ defined by $T^*(f) = f \circ T$ is a map of algebras. The observation that $T^{-1}(V(f)) = V(f \circ T)$ for any $f \in A(W)$, together with the preceding remarks, implies that $T: V \to W$ is continuous.

3. The main theorem. Central to the proof of the main theorem is the following elementary lemma.

3.1 LEMMA. *Let D be an integral domain and $f(X_1,\cdots,X_n) \in D[X_1,\cdots,X_n]$. If $f(X_1,\cdots,a) = 0$ for infinitely many $a \in D$, then $f(X_1,\cdots,X_n) = 0$.*

Proof: If $n > 1$, then $D' = D[X_1,\cdots,X_{n-1}]$ is a domain and $D[X_1,\cdots,X_n] = D'[X_n]$. Thus we may assume $n = 1$. If k is the field of quotients of D, then $D[X_1] \subseteq k[X_1]$, so we may also assume $D = k$. But the proof is clear in this case.

An immediate consequence of 3.1 is:

3.2 *If k is an infinite field and $0 \neq f(X_1,\cdots,X_n) \in k[X_1,\cdots,X_n]$, then there are $a_1,\cdots, a_n \in k$ such that $f(a_1,\cdots,a_n) \neq 0$.*

Let $V = k^n$ and let $\{x_1,\cdots,x_n\}$ be the dual basis of the natural basis for k^n. If k is infinite, the (surjective) algebra map $\tau: k[X_1,\cdots,X_n] \to A(k^n)$ determined by $\tau(X_i) = x_i$ is an isomorphism by 3.2. In this case we identify the polynomial ring $k[X_1,\cdots,X_n]$ with $A(k^n)$. Let $|S|$ denote the cardinality of a set S. Now we are ready to prove the main theorem.

3.3 THEOREM. *Let V be a finite-dimensional vector space over an infinite field k. If $\mathscr{B}$ is a family of proper closed subsets of V such that $\cup\{B: B \in \mathscr{B}\} = V$, then $|\mathscr{B}| = |k|$.*

Proof: Without loss of generality we may assume $V = k^n$. If $B \in \mathscr{B}$, then $B \subset V(f)$ for some $0 \neq f \in A(k^n)$; therefore we may assume $\mathscr{B} = \{V(f): f \in \mathscr{F}\}$, where $\mathscr{F}$ is a set of non-zero functions of $A(k^n) = k[X_1,\cdots,X_n]$.

If the theorem is false, let n be least integer for which it is false. Then there is a family of non-zero polynomials $\mathscr{F} \subseteq k[X_1,\cdots,X_n]$, satisfying $|\mathscr{F}| < |k|$ and $\cup\{V(f): f \in \mathscr{F}\} = k^n$. For $f = f(X_1,\cdots,X_n) \in \mathscr{F}$ and $a \in k$, let $f_a = f(X_1,\cdots,a)$. By 3.1 the set $\{a \in k: f_a = 0\}$ is finite (or void). Since k is infinite, $|S| < |k|$, where $S = \{a \in k: f_a = 0 \text{ some } f \in \mathscr{F}\}$. Choose $\alpha \in k \backslash S$. Then $V(f_\alpha)$ is a proper subset of k^{n-1} by 3.2, and $\cup\{V(f_\alpha): f \in \mathscr{F}\} = k^{n-1}$ by assumption. This contradicts the minimality of n.

3.3 implies that a finite-dimensional vector space over an infinite field is not the union of a family $\mathscr{B}$ of proper subspaces if $|\mathscr{B}| < |k|$.

The following corollary shows that the union of a family of proper closed sets is "thin" if $|\mathscr{B}| < |k|$.

3.4 COROLLARY. *Let V be a finite-dimensional vector space over an infinite field k, $\mathscr{B}$ a family of proper closed subsets of V. If U is a(non-void) open subset of V and $U \subseteq \cup\{B: B \in \mathscr{B}\}$, then $|\mathscr{B}| = |k|$.*

Proof: Let $\mathscr{B}' = \mathscr{B} \cup \{V \setminus U\}$. Then $\mathscr{B}'$ is a family of proper closed sets, and $V = \cup\{B: B \in \mathscr{B}'\}$ by assumption. By 3.3, $|\mathscr{B}| = |k|$.

3.5 COROLLARY. *Suppose V is a finite-dimensional vector space over an infinite field k, $\mathscr{F}$ a subset of $A(V)$ satisfying $|\mathscr{F}| < |k|$. If U is any (non-void) open subset of V, then there is a $u \in U$ such that $f(u) \neq g(u)$ for all distinct $f, g \in \mathscr{F}$.*

Proof: Let $\mathscr{F}' = \{f - g: f, g \in \mathscr{F} \text{ distinct}\}$. Then $0 \notin \mathscr{F}'$ and $|\mathscr{F}'| < |k|$. Since $V(f')$ is proper for all $f' \in \mathscr{F}'$, we conclude from 3.4 that $U \nsubseteq \cup\{V(f'): f' \in \mathscr{F}'\}$. So choose $u \in U$ such that $u \notin V(f')$ all $f' \in \mathscr{F}'$.

One should note that 3.5 gives a general condition under which functional may be distinguished by a single vector. Identifying V with V^{**} in the finite-dimensional case we have as a consequence of 3.5:

3.6 COROLLARY. *Suppose V is a finite-dimensional vector space over an infinite field k, S a subset of V such that $|S| < |k|$. If U is any (non-void) open set of V^*, then there is an $f \in U$ such that $f(s) \neq f(t)$ for all distinct $s, t \in S$.*

One easy consequence of 3.6 is that if S is a subset of an n-dimensional vector space V over an infinite field k satisfying $0 \notin S$ and $|S| < |k|$, then there is an $n-1$ dimensional subspace W of V such that $W \cap S = \emptyset$. (Choose an appropriate $f \in U = V^*$and let $W = \ker f$).

We conclude with a result about the action of open sets of endomorphisms of V on closed subsets of V.

3.7 PROPOSITION. *Let V be a finite-dimensional vector space over an infinite field k, B a proper closed subset of V, and U any (non-void) open subset of $\mathrm{End}_k V$. Suppose S is a subset of V such that $0 \notin S$ and $|S| < |k|$. Then there is a $u \in U$ such that $u^{-1}(B) \cap S = \emptyset$.*

Proof: Since $B \subset V(f)$ for some $0 \neq f \in A(V)$ we may assume $B = V(f)$. For each $s \in S$ let π_s: $\mathrm{End}_k V \to V$ be defined by $\pi_s(T) = T(s)$. Then π_s is surjective since $s \neq 0$. Since $V(f)$ is proper and π_s continuous, $B_s = \pi_s^{-1}(V(f))$ is a proper closed subset of $\mathrm{End}_k V$ all $s \in S$. By 3.4 $U \nsubseteq \cup \{B_s: s \in S\}$. So choose $u \in U$ such that $u \notin B_s$ all $s \in S$. Thus $\pi_s(u) = u(s) \notin V(f)$ which implies $s \notin V(f \circ u) = u^{-1}(V(f))$ all $s \in S$.

Reference

1. I. N. Herstein, Topics in Algebra, Blaisdell, Waltham, Mass., 1966.

ON A PROBLEM OF GOLOMB ON POWERFUL NUMBERS

ANDRZEJ MAKOWSKI, University of Warsaw

S. W. Golomb [1] defined a powerful number as a positive integer which for every prime number p is divisible by p^2 provided it is divisible by p. He asked whether there exist positive integers $\neq 1$ and 4 which are in infinitely many ways representable as the differences of two relatively prime powerful numbers.

We prove below that the answer to this question is in the affirmative.

It is known [2], p. 56 that every prime number $p \equiv 1 \pmod 8$ is representable in the form $x^2 - 2y^2$ and in view of the identity

$$x^2 - 2y^2 = (3x + 4y)^2 - 2(2x + 3y)^2$$

there are infinitely many such representations. Evidently, in every such representation x is odd and y is even, hence $p = x^2 - 8z^2$ and both x^2 and $8z^2$ are relatively prime powerful numbers. Because there are infinitely many prime numbers $\equiv 1 \pmod 8$ we infer that there are infinitely many numbers satisfying the Golomb's conditions.

References

1. S. W. Golomb, Powerful numbers, this MONTHLY, 77 (1970) 848–852.
2. B. A. Venkov, Elementary number theory, Wolters-Noordhof, Groningen 1970.

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

DOES THERE EXIST MORE THAN ONE BANACH *-ALGEBRA WITH DISCONTINUOUS INVOLUTION?

R. S. DORAN, Texas Christian University

A ***-algebra** is a complex associative linear algebra A with a mapping $x \to x^*$ of A into itself such that for $x, y \in A$ and complex λ: (a) $(x+y)^* = x^* + y^*$; (b) $(xy)^* = y^*x^*$; (c) $(\lambda x)^* = \bar{\lambda} x^*$ ($\bar{\lambda}$ is the complex conjugate of λ); and (d) $x^{**} = x$. The map $x \to x^*$ is called an **involution**; because of (d) it is clearly bijective. An algebra which is also a Banach space satisfying $\| xy \| \leqq \| x \| \cdot \| y \|$ for all x, y is called a **Banach algebra**. A Banach algebra which is also a *-algebra is called a **Banach *-algebra.**

Typical examples of Banach *-algebras are the complex numbers C with the usual multiplication, involution $\lambda^* = \bar{\lambda}$ (complex conjugation), and absolute value norm; the algebra $C(X)$ of bounded continuous complex-valued functions on a topological space X with pointwise multiplication $(fg)(t) = f(t)g(t)$, involution $f^*(t) = \overline{f(t)}$, and sup norm; the algebra $\mathscr{B}(\mathscr{H})$ of bounded linear operators on a Hilbert space $\mathscr{H}$ with composition as multiplication, involution $T \to T^*$ (the adjoint of T), and operator norm; the group algebra $L^1(G)$ of a locally compact abelian group G with multiplication

$$(f*g)(t) = \int_G f(t-s)g(s)\,ds,$$

involution $f^*(t) = \overline{f(-t)}$, and L^1-norm; and the algebra $A(D)$ of continuous complex-valued functions on the closed unit disc D which are analytic on the interior of D with pointwise multiplication, involution $f^*(\lambda) = \overline{f(\bar{\lambda})}$, and sup norm.

For particular topological spaces X, variations of the involution in the second example can be given. For instance, if $X = [0,1]$ with the usual topology, then $f^*(t) = \overline{f(1-t)}$ defines an involution in $C(X)$. As a second illustration, let $X = [0,1] \cup \{2,3\}$ with the usual relative topology of the reals and, for $f \in C(X)$, define $f^*(t) = \overline{f(t)}$ if $t \in [0,1]$, $f^*(2) = \overline{f(3)}$, and $f^*(3) = \overline{f(2)}$.

To see how extensive the class of Banach *-algebras are we note that *every* Banach algebra A can be isometrically embedded as a closed two-sided ideal of a Banach *-algebra B. Indeed, let $B = A \times A$ and define

$$(x, y) + (w, z) = (x + w, y + z),\ (x, y)(w, z) = (xw, yz),$$
$$\lambda(x, y) = (\lambda x, \bar{\lambda} y),\quad (x, y)^* = (y, x),$$
$$\| (x, y) \| = \max\{\| x \|, \| y \|\}.$$

Then B is a Banach *-algebra and the map $x \to (x, 0)$ is an isometric embedding of A in B.

The involution in a Banach *-algebra A is said to be **continuous** if there exists a constant M such that $\| x^* \| \leqq M \| x \|$ for all $x \in A$; if no such constant exists the involution is said to be **discontinuous**. All of the involutions described above are continuous. On the other hand, Banach *-algebras with discontinuous involution do not appear to be numerous. In fact, the following example due to F. F. Bonsall (see [**2**] p. 704), is the only one known to the author.

Let A be an infinite-dimensional Banach space over the complex numbers, and make A into a Banach algebra by giving it the trivial multiplication, i.e., $ab = 0$ for $a, b \in A$. Let E be a Hamel basis for A, chosen so that $\| x \| = 1$ for each $x \in E$. Let $\{x_n\}$ be a sequence of distinct elements of E and define x_n^* by

$$x_{2n-1}^* = n x_{2n},\quad x_{2n}^* = \frac{1}{n} x_{2n-1}\quad (n = 1, 2, \cdots).$$

For all other elements of E, let $x^* = x$, and then extend the mapping $x \to x^*$ to all of A by conjugate linearity; that is,

$$(\lambda_1 y_1 + \cdots + \lambda_k y_k)^* = \bar{\lambda}_1 y_1^* + \cdots + \bar{\lambda}_k y_k^* \qquad (\lambda_i \in C, y_i \in E).$$

Then $x \to x^*$ is an involution on A which is clearly not continuous since $\| x_{2n} \| = 1$ and $\| x_{2n}^* \| = 1/n$.

The example just described is not particularly satisfying because the multiplication in the algebra is trivial. Of course, by adjoining an identity to A in the usual way (form $A \times C$ with coordinatewise linear operations, and define $(x, \lambda)(y, \mu) = (xy + \lambda y + \mu x, \lambda\mu)$, $(x, \lambda)^* = (x^*, \bar{\lambda})$, and $\| (x, \lambda) \| = \| x \| + |\lambda|$) one obtains a Banach *-algebra with identity and nontrivial multiplication which has discontinuous involution. However, this is hardly more satisfying than the example itself.

PROBLEM. *Find an example of a Banach *-algebra with discontinuous involution which has nontrivial multiplication and which is not obtained from an algebra with trivial multiplication by adjoining an identity.*

The problem requires that the reader construct an algebra distinct from the one above. In view of several recent papers concerning removal of continuity from the involution (e.g., [**3**], [**5**], [**8**]), a solution to the problem would be of considerable interest. Since the norm in any *semi-simple* Banach algebra is unique up to equivalence [**4**], it follows easily that every involution on such an algebra is automatically continuous (simply note that $\| x \|_0 = \| x^* \|$ defines a second complete algebra norm).

The reader must restrict his attention accordingly when looking for examples.

A subalgebra B of a *-algebra is called a ***-subalgebra** if $x \in B$ implies that $x^* \in B$. The involution in a Banach *-algebra A is said to be **locally continuous** if it is continuous when restricted to each maximal commutative *-subalgebra of A. Every continuous involution is clearly locally continuous. What about the converse?

CONJECTURE. *If the involution in a Banch *-algebra is locally continuous, then it is continuous.*

An affirmative solution to this conjecture would indeed be interesting! On the other hand, a counter-example to the conjecture would again require that the reader construct a new example of a Banach *-algebra with discontinuous involution. An excellent elementary introduction (suitable for advanced undergraduates) to Banach algebras can be found in [**9**], and much more extensive treatments are given in [**6**] and [**7**].

References

1. P. Civin and B. Yood, Involutions on Banach algebras, Pacific J. Math., 9(1959) 415–436.

2. J. Duncan, The continuity of the involution on Banach *-algebras, J. London Math. Soc., 41(1966) 701–706.

3. J. W. M. Ford, A square root lemma for Banach *-algebras, J. London Math. Soc., 42(1967) 521–522.

4. B. E. Johnson, The uniqueness of the (complete) norm topology, Bull. Amer. Math. Soc., 73(1967) 537–539.

5. I. S. Murphy, Continuity of positive linear functionals on Banach *-algebras, Bull. London Math. Soc., 1(1969) 171–173.

6. M. A. Naimark, Normed Rings, 2nd ed., Noordhoff, Groningen, the Netherlands, 1964.

7. C. E. Rickart, General Theory of Banach Algebras, Van Nostrand, Princeton, N.J., 1960.

8. S. Shirali, Representability of positive functionals, J. London Math. Soc., (2), 3(1971) 145–150.

9. G. F. Simmons, Introduction to Topology and Modern Analysis, McGraw-Hill, New York, 1963.

HOW SEPARABLE IS A SPACE?

ALBERT WILANSKY, Lehigh University

In 1944 it was proved that the product P of c (or fewer) separable spaces X_α is separable. For historical discussion and proof see [**1**]. Thus P has a countable dense subset D. However, [**2**], D need not be sequentially dense even if each X_α has a sequentially dense countable subset. This raises the problem of finding conditions under which a space has a countable sequentially dense subset. More generally, we make the following conjecture.

For an infinite cardinal m, say that a set D is ***m*-dense in X** if there exists a totally

ordered set O with m or fewer members such that each $x \in X$ is the limit of a net in D defined on O.

CONJECTURE. *Let m be an infinite cardinal and* $\{X_\alpha : \alpha \in A\}$ *a collection of spaces each of which has an m-dense subset with m points (or less). Then* πX_α *has an m-dense subset with m points (or less) if* $|A| < 2^m$, *and need not if* $|A| = 2^m$.

For ordinary density we have that πX_α has a dense subset with m points (or less) if $|A| \leqq 2^m$. See [**1**].

References

1. W. W. Comfort, A short proof of Marczewski's separability theorem, this MONTHLY, 76 (1969) 1041–1042.

2. N. J. Kalton, A barreled space without a basis, Proc. Amer. Math. Soc., 26(1970) 465–466.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

A NOTE ON EXT AND TOR

JERRY HOPPONEN, University of Colorado

A basic course in homological algebra usually develops a number of facts about the behavior of certain functors with respect to sums and products. However, when considering sums and products over an infinite indexing set, the isomorphisms

$$\text{Tor}_n(\textstyle\sum C_\alpha, A) \simeq \sum \text{Tor}_n(C_\alpha, A) \tag{1}$$

$$\text{Ext}^n(\textstyle\sum C_\alpha, A) \simeq \prod \text{Ext}^n(C_\alpha, A) \tag{2}$$

$$\text{Ext}^n(C, \textstyle\prod A_\alpha) \simeq \prod \text{Ext}^n(C, A_\alpha) \tag{3}$$

are often omitted, or proved by induction on n, as in Rotman [**2**]. In this note we establish these results using only basic facts from homological algebra.

Recall that by definition

$$\text{Tor}_n(C, A) = H_n(X \otimes A),$$

$$\text{Ext}^n(C, A) = H_n[\text{Hom}(X, A)] = H_n[\text{Hom}(C, Y)],$$

where $X \to C$ is a projective resolution and $A \to Y$ is an injective resolution.

Let H denote the homology functor. We shall rely on the following standard results:

THEOREM A. *If* $\{X_\alpha\}$ *is a collection of differential R-modules, then*

$$H(\sum X_\alpha) \simeq \sum H(X_\alpha) \text{ and } H(\prod X_\alpha) \simeq \prod H(X_\alpha).$$

Proof. Left as an easy exercise.

THEOREM B. *If* $\{A_\alpha\}$ *and* $\{B_\alpha\}$ *are collections of R-modules and A and B are R-modules, then*

$$\mathrm{Hom}(\sum A_\alpha, B) \simeq \prod \mathrm{Hom}(A_\alpha, B),\ \mathrm{Hom}(A, \prod B_\alpha) \simeq \prod \mathrm{Hom}(A, B_\alpha),$$

and

$$(\sum A_\alpha) \otimes B \simeq \sum (A_\alpha \otimes B).$$

Proof. See [**2**] and [**1**].

Proof of isomorphism 1. Let $\{X_n^\alpha, \partial^\alpha\}_n$ be a projective resolution of C_α, for each α, and let $X_n = \sum_\alpha X_n^\alpha$. Note that X_n is projective. The exactness of ∂^α and the universal property of direct sums induce an exact differential $\partial: X_n \to X_{n-1}$ so that $X = \{X_n, \partial\}$ becomes a projective resolution of $\sum C_\alpha$. Therefore

$$\begin{aligned}\mathrm{Tor}_n(\sum C_\alpha, A) = H[(\sum X_n^\alpha) \otimes A] &\simeq H[\sum (X_n^\alpha \otimes A)] \\ &\simeq \sum H(X_n^\alpha \otimes A) = \sum \mathrm{Tor}_n(C_\alpha, A).\end{aligned}$$

Isomorphism 2 follows similarly.

Proof of isomorphism 3. Dualize the necessary parts above to get an injective resolution $Y = \{\prod_\alpha Y_\alpha^n, \delta\}^n$ of $\prod A_\alpha$. As before,

$$\begin{aligned}\mathrm{Ext}^n(C, \prod A_\alpha) = H[\mathrm{Hom}(C, \prod Y_\alpha^n)] &\simeq H[\prod \mathrm{Hom}(C, Y_\alpha^n)] \\ \simeq \prod H[\mathrm{Hom}(C, Y_\alpha^n)] &= \prod \mathrm{Ext}^n(C, A_\alpha).\end{aligned}$$

References

1. S. MacLane, Homology, Academic Press, New York, 1963, p. 141.
2. J. Rotman, Notes on Homological Algebra, Van Nostrand, Princeton, N. J., Math. Studies, 1970, pp. 15–18, pp. 145–146.

AN HISTORICAL NOTE ON THE PARITY OF PERMUTATIONS

T. L. BARTLOW, Villanova University

Every beginning algebra student learns that the number of transpositions into which a given permutation can be decomposed is either always even or always odd. Many students find the traditional proof, involving the function

$$P = (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1)(x_3 - x_2) \cdots (x_n - x_{n-1}),$$

unsatisfactory, because, as Herstein remarks, the polynomial "seems extraneous to the matter at hand" [**7**, p. 67]. Many alternative proofs have been offered, so many that we wonder why the traditional proof maintains its place in textbooks [**1**, **5**, **6**, **8** (p. 36), **10**, **11**, **12**, **14**]. Here we offer two more alternatives which derive from the origins of the subject.

I. The first is to explain that early studies of permutations occurred in a context in which the polynomial P is quite natural. Mathematicians of the sixteenth century knew that the coefficients of a polynomial could be expressed as elementary symmetric functions of the roots of the polynomial. In 1770–1771, Lagrange [**9**] and Vandermonde [**13**] made the first efforts to exploit this fact to discuss the question of solvability of polynomials of degree greater than four. They recognized that any formula solving a general polynomial of degree n in terms of the coefficients of the polynomial must be a symmetric function in the n roots. This realization suggested to them the importance of studying the effect of permutations on functions of n variables. In 1815 Cauchy published the results of a careful study of this question [**2**, **3**]. Cauchy credits Vandermonde with observing that the function P is a typical example of an alternating function, although Vandermonde appears to have made this observation only for $n=3$. Moreover, Cauchy proves that every alternating function of n variables is divisible by P. The first extensive study of permutations and permutation groups appeared in 1844–45 in several papers by Cauchy.

Thus early work on permutation groups was largely motivated and informed by investigations of the effect of permutations on a function of several variables, investigations in which the function P had a prominent role. An old text on Galois Theory takes this point of view [**4**].

II. The other alternative is to use what appears to be the original proof of the theorem in question, which did not involve P. In [**3**, pp. 98–104] Cauchy gives a proof which relates the parity of a permutation to the number of cycles which it involves.

LEMMA. *Every permutation is uniquely a product of disjoint cycles.*

Cauchy's proof is the one in use today. It will be important to count the number of cycles in a given permutation. For this purpose Cauchy counts one-cycles. Thus the identity permutation is, in modern notation, $(1)(2)\cdots(n)$ and $(1,3,4)$ is properly designated by $(1,3,4)(2)\cdots(n)$.

LEMMA. *If a product of g disjoint cycles is multiplied by a transposition, the result involves $g \pm 1$ disjoint cycles.*

Proof. Let α be a permutation involving g cycles, and let (a,b) be a transposition. If a and b belong to the same cycle of α, the product has the form

$$(a,c,\cdots,d,b,f,\cdots,h)\cdots(a,b)=(a,c,\cdots,d)(b,f,\cdots,h)\cdots,$$

which contains $g+1$ cycles. If a and b are in different cycles of α we have

$$(a, c, \cdots, f)(b, d, \cdots, h) \cdots (a, b) = (a, c, \cdots, f, b, d, \cdots, h) \cdots,$$

which involves $g-1$ cycles. The argument applies even if a or b stands alone in a cycle.

LEMMA. *Let the permutations of S_n be partitioned into classes according as they involve an even or an odd number of cycles. If a permutation is multiplied by a sequence of transpositions, it does or does not change classes according as the number of transpositions is odd or even.*

Proof. This is immediate from the preceding lemma.

THEOREM. *No permutation can be the product both of an even and of an odd number of permutations.*

Proof. Let α be any permutation and consider the partition of the preceding lemma. Let the identity permutation be multiplied by α, regarded as a product of transpositions. The number of transpositions is even if and only if α is in the same class as the identity and odd if any only if α is in the other class. Of course, α cannot be in both classes.

It follows that the partition of the preceding lemma is, in fact, the partition of S_n into even and odd permutations. Since the identity permutation involves n cycles even permutations involve an even number of cycles if and only if n is even. Thus, if $c(\pi)$ is the number of cycles involved in π, then π is even if and only if $n-c(\pi)$ is even. This observation of Cauchy's is the starting point for Phillips [**11**].

References

1. J. L. Brenner, A new proof that no permutation is both even and odd, this MONTHLY, 64 (1957) 499–500.

2. A. -L. Cauchy, Mémoire sur le nombre des valeurs qu'une fonction peut acquérir lorsq'on y permute de toutes les manières possibles les quantités qu'elle renferme, J. l'École Polytechnique, 10 (1815) 1–28; Oeuvres Complètes, ser. II, vol. 1, pp. 64–90.

3. A. -L. Cauchy, Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opérées entre les variables qu'elles renferment, J. de l'École Polytechnique, 10 (1815) 29–112; Oeuvres Complètes, ser II. vol. pp. 91–169.

4. E. Dehn, Algebraic Equations: An Introduction to the Theories of Lagrange and Galois, Dover, New York, 1960.

5. E. L. Gray, An alternative proof for the invariance of parity of a permutation written as a product of transpositions, this MONTHLY, 70 (1963) 995.

6. I. Halperin, Odd and even permutations, Canadian Math. Bull., 3 (1960) 185–186.

7. I. N. Herstein, Topics in Algebra, Blaisdell, Waltham, Mass., 1964.

8. N. Jacobson, Lectures in Abstract Algebra, vol. I. Van Nostrand, Princeton, N. J., 1951.

9. J. L. Lagrange, Réflexions sur la résolution algébrique des équations, Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Berlin (1770–1771); Oeuvres de Lagrange 3, 205–421.

10. Hans Liebeck, Even and odd permutations, this MONTHLY, 76 (1969) 668.

11. W. Phillips, On the definition of even and odd permutations, this MONTHLY, 74 (1967) 1249–1251.

12. E. L. Spitznagel, Note on the alternating group, this MONTHLY, 75 (1968) 68–69.

13. A. Vandermonde, Mémoire sur la résolution des équations, Histoire de l'Académie Royale des Sciences, 88 (1771) 365–416.

14. C. E. Weil, Another approach to the alternating subgroup of the symmetric group, this MONTHLY, 71 (1964) 545–546.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

Notice

The December, 1971 issue of the *American Scientist* contains an interesting article by René Thom entitled "'Modern' Mathematics: An Educational and Philosophical Error?" Since this journal is widely circulated and, we hope, easily accessible to the readers of the MONTHLY, the editors of this Section have decided to recommend this article to you instead of reprinting it.

REPORT OF THE COMMITTEE ON THE UNDERGRADUATE PROGRAM IN MATHEMATICS, JANUARY, 1972

Commission. Following the directives of the National Science Foundation, no blanket proposal for the biennium 1972–74 has been submitted. However, the following six proposals for separate projects were prepared, approved by the Executive and Finance Committees and submitted on November 1, 1971, for consideration by the Foundation.

(1) A Proposal to Produce Case Studies and Resource Materials for the Teaching of Applied Mathematics at the Advanced Undergraduate Level.

(2) A Proposal for Improving the Teaching of Mathematics in the Technical and Occupational Programs of Two Year Colleges.

(3) A Proposal for a Survey of Innovative Methods of Teaching Undergraduate Mathematics and the Dissemination of the Findings.

(4) A Proposal for the Support of Speakers to Discuss CUPM Reports at Professional Meetings.

(5) A Proposal to Hold Conferences on Selected CUPM Reports.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before December 31, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2367. *Proposed by Erwin Just, Bronx Community College*

Let F_n be the nth term of the sequence defined by

$$F_n = -F_{n-1} - 2 \cdot F_{n-2},\ F_1 = 1,\ F_2 = -1.$$

Prove that $2^{n+1} - 7F_{n-1}^2$ is a perfect square.

E 2368. *Proposed by C. V. Heuer, Concordia College*

Prove that if $1 < x_1 < x_2 < \cdots < x_k < y_1 < y_2 < \cdots < y_m$ are integers such that $\Sigma x_i \geqq \Sigma y_i$, then $\prod x_i > \prod y_i$.

E 2369. *Proposed by Harry Lass, Jet Propulsion Laboratory, California Institute of Technology*

For the two-dimensional symmetric random walk starting at the origin, show that the probability of reaching the point (1,0) before reaching any other point on the line $x = 1$, is $1 - 2/\pi$.

E 2370. *Proposed by John Hyde, Student, St. Olaf College.*

Let R be a ring with identity, and let $R[x]$ be the ring of polynomials over R in the indeterminate x. A current modern algebra textbook asks the student to prove that $R[x]$ cannot contain $\sqrt{x}$; that is, $R[x]$ cannot contain a polynomial $f(x)$ such that $[f(x)]^2 = x$. Find an example of a ring R and a polynomial $f(x)$ that disproves this. Can R be commutative?

E 2371. *Proposed by M. H. Greenblat*

A clever graduate student (CGS) was discussing a mathematical problem with his friend, the absent-minded professor (AMP). The CGS asked, "Do you remember that cubic equation we solved several weeks ago, you know the one in which the coefficient of all the terms were positive integers? It had integral roots, and the coefficient of the cubic term was unity?"

AMP – "Well, I remember it only vaguely."
CGS – "I'd like to reconstruct it. Do you remember the value of the constant term?"
AMP – "Not precisely. I remember it was either 2450 or 2540."
CGS – "Well, do you remember the coefficient of the square term?"
AMP – "I'm afraid not, but it wouldn't help you even if I did remember it." (In this, he underestimated the CGS.)

CGS — "Aha! Was the coefficient of the linear term as high as it could possibly be?"
AMP — "Yes."

At this point, the CGS knew the equation in question. You can, too, with the above information.

E 2372. *Proposed by E. T. Wang, University of British Columbia*

Let A be an $n \times n$ matrix with entries zero and one, such that each row and each column contains precisely k ones. A *generalized diagonal* of A is a set of n elements of A such that no two elements appear in the same row or the same column. Show that A has at least k pairwise disjoint generalized diagonals, each of which consists entirely of ones.

SOLUTIONS OF ELEMENTARY PROBLEMS

Sequences with Precisely $k+1$ k-Blocks

E 2307 [1971, 792]. *Proposed by D. M. Bloom, Brooklyn College*

Given an infinite sequence, by a *k-block* in the sequence we mean a block of k consecutive terms. Prove or disprove: there exists an infinite sequence S such that (a) for all n, $S_n = 0$ or 1; (b) for every k, the sequence S contains exactly $k+1$ different k-blocks. (Note: a different problem involving the number of k-blocks in a sequence of zeros and ones appeared in the 1955 Putnam Examination.)

Solution by D. E. Knuth, Stanford University. If ρ is any irrational number between 0 and 1, then we obtain such a sequence by setting $S_n = [(n+1)\rho] - [n\rho]$, where the square brackets indicate the usual greatest integer function.

To prove this, let (x) denote the fractional part of the real number x: $(x) = x - [x]$. For any fixed positive integer k, take the k numbers $(-\rho), (-2\rho), \cdots, (-k\rho)$, and arrange them in increasing order: $a_1 < a_2 < \cdots < a_k$. Then define $a_0 = 0$ and $a_{k+1} = 1$. Whenever $(n\rho)$ lies between a_j and a_{j+1}, the k-block $S_n S_{n+1} \cdots S_{n+k-1}$ has a fixed value B_j, since the points $x, x+\rho, x+2\rho, \cdots, x+k\rho$ do not pass any integer values as x varies from a_j to a_{j+1}. It follows that there are at most $k+1$ different k-blocks. Moreover, if we consider what happens when x varies past a_{j+1}, we see that B_{j+1} is formed from B_j by changing an adjacent pair of elements from $\cdots 01 \cdots$ to $\cdots 10 \cdots$ or by changing the final element from 0 to 1. Thus, if we regard the B_j as binary numbers, we have $B_0 < B_1 < \cdots < B_k$. Since the set of all $(n\rho)$ is dense in the interval $[0,1]$, all $k+1$ of these distinct k-blocks must occur.

Perhaps the simplest sequence of the required type is what I called the "Fibonacci string sequence" in my book, *The Art of Computer Programming*, Vol. 1, Addison-Wesley, 1968, Exercise 1.2.8–36. Define $Q_0 = 0$, $Q_1 = 1$, and $Q_{n+2} = Q_{n+1}Q_n$,

the operation being concatenation. Thus $Q_2 = 10$, $Q_3 = 101$, $Q_4 = 10110$, etc. The limiting sequence has Bloom's property. I note in my book (p. 493) that the Fibanocci string sequence is the special case of my general construction above, which is obtained by taking $\rho = \frac{1}{2}(\sqrt{5} - 1)$, the reciprocal of the "golden mean."

Also solved by L. J. Guibas, Harry Lass, O. P. Lossers (Netherlands), J. G. Mauldon, P. L. Montgomery, and the proposer.

Editor's Comment. Several solutions were submitted which considered only doubly infinite "sequences." In this case, a sequence such as ...0001000... provides a solution. G. A. Hedlund notes that related problems have been considered in Marston Morse and G. A. Hedlund, *Symbolic dynamics II, Sturmian trajectories*, Amer. J. Math. 62 (1940), 1–42, and Benjamin G. Klein, *Homomorphisms of symbolic dynamical systems*, to appear in Math. Syst. Theory.

Are All Weird Numbers Even?

E 2308 [1971, 792]. *Proposed by Stan Benkoski, Pennsylvania State University*

Call the natural number n *semiperfect* if there is a collection of distinct proper divisors of n whose sum is n. In order that n be semiperfect it is necessary that it be perfect or abundant. (A natural number n is perfect (abundant) if the sum of the proper divisors of n is equal to (greater than) n.)

(a) Show that the condition is not sufficient. (b) Are all abundant numbers semiperfect?

Comment by the proposer. What I have called semiperfect numbers have been studied by W. Sierpinski, *Sur les nombres pseudoparfaits*, Mat. Vesnik. 2 (1965), 212–213. I shall call a number which is abundant but not semiperfect a *weird* number; the only weird numbers not exceeding 10,000 are the following: 70, 836, 4030, 5830, 7192, 7912, and 9272. The question of whether there exist odd weird numbers may be very difficult. Professor Paul Erdös has offered $10 for the first example of an odd weird number, and $25 for the first proof that none can exist.

A number is *primitive semiperfect* if it is semiperfect, but it is not divisible by any other semiperfect number. There are infinitely many primitive semiperfect numbers. There are also infinitely many weird numbers and, in fact, the set of weird numbers has positive density.

Editor's comment. If n is semiperfect and if $d \mid n$, write $d' = n/d$. Then $n = \Sigma d$ implies that $1 = \Sigma 1/d'$, so that 1 is expressed as a sum of Egyptian fractions. This has interest when n is odd.

Professor Erdös has also offered $25 for a solution to the following related question: For every $2 \leq c < \infty$, is there an integer n which is not semiperfect but which satisfies $\sigma(n)/n > c$? That is, is $\sigma(n)/n$ bounded as n ranges through the set of weird numbers?

The fact that 70 is a weird number was noted by Lew Kowarski, Harry Lass, and the St. Olaf College Students.

Distinct Representatives for a Collection of Finite Sets

E 2309 [1971, 792]. *Proposed by Václav Chvátal, University of Waterloo, Ontario*

Prove the following: Let $A_1, \cdots, A_n$ be finite sets. If

$$\sum_{1 \leq i < j \leq n} |A_i \cap A_j| \frac{|A_1| \cdots |A_n|}{|A_i| \cdot |A_j|} < 1,$$

then the sets $A_1, \cdots, A_n$ have a system of distinct representatives (i.e., there are $a_1, a_2, \cdots, a_n$ such that $a_i \in A_i$ and $a_i \neq a_j$ for $i \neq j$).

Solution by the proposer. There are exactly $|A_1| \cdot |A_2| \cdots |A_n|$ mappings

$$f\colon \{1, 2, \cdots, n\} \to \bigcup_{i=1}^{n} A_i \tag{1}$$

such that $f(i) \in A_i$ for $i = 1, 2, \cdots, n$. The problem asks if there exists a function of type (1) which is one-to-one. If a function h is of type (1) and is not one-to-one, then there are at least two distinct integers $i, j \in \{1, 2, \cdots, n\}$ such that $h(i) = h(j)$. The number of mappings of type (1) which are not one-to-one then does not exceed

$$\begin{aligned} &\sum_{1 \leq i < j \leq n} |A_i \cap A_j| \frac{|A_1| \cdot |A_2| \cdots |A_n|}{|A_i| \cdot |A_j|} \\ &= |A_1| \cdot |A_2| \cdots |A_n| \sum_{1 \leq i < j \leq n} \frac{|A_i \cap A_j|}{|A_i| \cdot |A_j|} \\ &< |A_1| \cdot |A_2| \cdots |A_n| \quad \text{(by hypothesis).} \end{aligned}$$

Therefore, the number of functions of type (1) which are not one-to-one is strictly less than the number of functions of type (1) and hence the conclusion follows.

Also solved by the Bennett College Team, D. M. Bloom, Bobby Chapuis & C. C. Rousseau, John Christopher, M. G. Greening (Australia), David Kelly, Harry Lass, Robert Patenaude, David Sumner, and J. H. Timmermans (Netherlands).

A Categorical Impossibility

E 2310 [1971, 793]. *Proposed by Hal Forsey, San Francisco State College*

Does there exist a positive function f such that if x is rational and y is irrational, then $f(x)f(y) \leq |x - y|$?

I. *Solution by Simeon Reich, Israel Institute of Technology, Haifa.* The answer is no. Let R, I, and Q denote the reals, irrationals, and rationals respectively, and

suppose that the desired f exists. We note first that if $\{r_n\}$ is a sequence of rationals which converges to an irrational, then $f(r_n) \to 0$, and likewise if $\{y_n\}$ is a sequence of irrationals which converges to a rational, then $f(y_n) \to 0$. Now let $g: R \to R$ agree with f on I and vanish on Q. Then Q is the set of points where g is continuous, so that it must be a G_δ set. But it is not, by the Baire Category Theorem.

Alternatively, we can let $h: R \to R$ agree with f on Q and vanish on I. Since h is Lipschitzian on I and discontinuous on Q, I must be a set of the first category by Theorem 4 of G. A. Heuer, *A property of functions discontinuous on a dense set*, this MONTHLY, 73 (1966), 378–379. But it is not, again by the Baire Category Theorem.

II. *Solution by Pavel Kostyrko, Bratislava, Czechoslovakia.* We characterize such functions in the following theorem.

THEOREM. *Let (M,d) be a metric space, and suppose that $X \subseteq M$ and $Y = M \setminus X$. Then there exists a (strictly) positive function $f: M \to R$ such that*

$$f(x)f(y) \leqq d(x,y) \text{ for all } x \in X,\ y \in Y \tag{1}$$

if and only if both X and Y are F_σ sets in M. [Note that if $X = \varnothing$ or $Y = \varnothing$, then (1) is vacuously satisfied—Ed.]

Proof. Suppose that such a function f exists and let $X_n = \{x \in X: f(x) \geqq 1/n\}$ for $n = 1, 2, \cdots$. We show that $\bar{X}_n \subseteq X$ for all n, where $\bar{Z}$ denotes the closure of Z in M. Suppose to the contrary that there exists a positive integer m and a y such that $y \in \bar{X}_m \setminus X$. Then $y \in Y$ and there exists a sequence $\{x_k\}$ of elements of X_m such that $x_k \to y$. Whence $f(y)/m \leqq f(x_k)f(y) \leqq d(x_k, y) \to 0$, implying that $f(y) = 0$, a contradiction. It follows that

$$X = \bigcup_{n=1}^{\infty} X_n \subseteq \bigcup_{n=1}^{\infty} \bar{X}_n \subseteq X,$$

so that X is an F_σ set in M. The proof for Y is analogous.

Conversely, suppose that X and Y are F_σ sets in M. Write $X = \bigcup_{n=1}^{\infty} F_n$ and $Y = \bigcup_{n=1}^{\infty} F_n^*$, where F_n and F_n^* are closed for $n = 1, 2, \cdots$, and where we assume without loss of generality that $F_1 \subseteq F_2 \subseteq \cdots$, and $F_1^* \subseteq F_2^* \subseteq \cdots$. The function f is defined as follows: If $x \in X$, let $n(x)$ denote the least positive n such that $x \in F_n$. Then define $f(x) = \min\{d(x, F_{n(x)}^*), 1\}$. If $y \in Y$, define $f(y)$ analogously. It can then be verified by checking cases that f has the required properties.

The problem is now solved by noting that the set of irrationals in R with the usual metric is not an F_σ set by the Baire Category Theorem.

III. *Solution by Charles Schelin, Wisconsin State University, La Crosse.* The answer is no. Suppose, to the contrary, that such a function exists. Let Q denote the set of rationals and H the set of irrationals. We note that if x is irrational and y is rational (or vice versa) then

$$f(x) \leqq \frac{1}{f(y)} |x - y|. \tag{*}$$

Let I_0 be any compact interval. Choose $x_1 \in H \cap I_0^o$, where I_0^o is the interior of I_0. By (*) we can find a neighborhood $N_1 = (x_1 - \delta_1, x_1 + \delta_1)$ of x_1 such that if $y \in Q \cap N_1$, then $f(y) < 1$. Now choose a fixed $y_1 \in Q \cap N_1 \cap I_0^o$; again by (*) we can find a neighborhood $M_1 = (y_1 - \eta_1, y_1 + \eta_1)$ of y_1 such that if $x \in H \cap M_1$, then $f(x) < 1$. Then for all $t \in M_1 \cap N_1 \cap I_0^o$ it is true that $f(t) < 1$. Select a nontrivial closed interval $I_1 \subseteq M_1 \cap N_1 \cap I_0^o \subseteq I_0$.

Continuing this process, we obtain a nested sequence $I_0 \supseteq I_1 \supseteq \cdots$ of closed bounded intervals with $f(t) < 1/n$ for all $t \in I_n$. By the Nested Interval Theorem, there is some $w \in \bigcap_{n=1}^{\infty} I_n$, forcing $f(w) < 1/n$ for every n; hence $f(w) \leqq 0$, a contradiction.

Also solved by Sheldon Axler, Bill Beckmann, Harold Donnelly, Neal Felsinger, Peter Frankl (Hungary), Gary Gunderson, G. A. Heuer, Terjéki József (Hungary), Peter Kuhfittig, Harry Lass, P. L. Montgomery, E. T. Ordman, Wolfe Snow, David Sumner, and the proposer.

Editorial Comment. Schelin's construction can be generalized as follows: Suppose that M is a compact metric space, and that Q is a subset of M such that both Q and its complement H are dense in every ball of positive radius. Then there cannot exist a strictly positive real-valued function f on M such that $f(x)f(y) \leqq d(x, y)$ for every $x \in Q$ and $y \in H$. Schelin's proof is interesting because it does not explicitly use the Baire Category Theorem.

The Compleat Cyclic Quadrilateral

E 2311 [1971, 793]. *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey*

Prove that, if a quadrilateral $A_1A_2A_3A_4$ can be inscribed in a circle, then the (six) lines drawn from the midpoints of A_pA_q perpendicular to A_rA_s (p, q, r, s distinct) are concurrent.

Solution by Sister Stephanie Sloyan, Georgian Court College, Lakewood, N.J. Assume that the circle is the unit circle and identify the point A_i with the complex number a_i in the usual manner. Then the line from the midpoint of the segment A_pA_q perpendicular to A_rA_s is given by

$$z - a_r a_s \bar{z} = \tfrac{1}{2}(a_p + a_q)\, \frac{a_p a_q - a_r a_s}{a_p a_q},$$

and it is easily calculated that all six lines pass through the point $\frac{1}{2}(a_1 + a_2 + a_3 + a_4)$. J. W. Clawson, *The complete quadrilateral*, Annals of Math. 20 (1918–1919), 232–261, calls this point the *orthic center* of the quadrilateral.

In a similar fashion one can show that the three lines joining the midpoint of A_pA_q to that of A_rA_s (p, q, r, s distinct) are each bisected by a point identified by Clawson as the *mean center* of the quadrilateral. Since the mean center is given by $\frac{1}{4}(a_1 + a_2 + a_3 + a_4)$, it follows that it lies halfway between the orthic center and the circumcenter.

Also solved by Michael Goldberg, Leonard Goldstone, M. G. Greening (Australia), N. G. Gunderson, V. F. Ivanoff, Lew Kowarski, Harry Lass, O. P. Lossers (Netherlands), Rick Troxel, and the proposer.

Editorial Note. This theorem and its solution appear on page 59 of Yaglom, *Complex Numbers in Geometry*, Academic Press, 1968, along with many other interesting properties of cyclic quadrilaterals, cyclic pentagons, etc. (see pages 54–68). The point of concurrence of this problem is called the *anticenter* by Lucien Droussent (*On a theorem of J. Griffiths*, this MONTHLY, 54 (1947), 538–540). The anticenter N is the midpoint of the quadrilateral's *Euler segment* which joins its circumcenter O to the center H of the circle through the four orthocenters H_m of the triangles $A_iA_jA_k$ ($\{i,j,k,m\} = \{1,2,3,4\}$); these orthocenters form a quadrilateral congruent to the given one and symmetric to it in point N. Furthermore, the eight points A_i and H_i lie by fours on four distinct pairs of circles, each pair having N as center of symmetry.

The eight congruent nine-point circles for the four triangles $A_iA_jA_k$ and four triangles $H_iH_jH_k$ all pass through N, and their centers lie on another congruent circle centered at N. Thus N can be called the *eight circle* point and this last circle the *eight point circle* for the quadrilateral.

There are four distinct Simson lines for the eight points A_m with triangles $A_iA_jA_k$ and H_m with triangles $H_i\,H_j\,H_k$, and these Simson lines all pass through N. In fact, one can form 280 (180 of which are distinct) pedal circles (and lines) by taking any one of these eight points with the triangle formed by any three others, and all of them pass through N.

The nine point centers N_m for the four triangles $A_iA_jA_k$ form a quadrilateral homothetic to $H_1H_2H_3H_4$ in center O with ratio $\frac{1}{2}$, hence homothetic to $A_1A_2A_3A_4$ in center G, 1/3 of the way from O to H, with ratio $-\frac{1}{2}$. Similarly, the nine-point centers N'_m for the triangles $H_iH_jH_k$ are homothetic to $H_1H_2H_3H_4$ in center G', 2/3 of the way from O to H, with ratio $-\frac{1}{2}$. Their common circumcircle has center N and radius half the given quadrilateral's circumradius. In a similar manner (see E 1740 [1965, 1026]) the centroids G_m for the triangles $A_iA_jA_k$ form a quadrilateral homothetic to $H_1H_2H_3H_4$ in center O with ratio 1/3, hence homothetic to $A_1A_2A_3A_4$ in center S (the mean center) 1/4 of the way from O to H, with ratio $-\frac{1}{3}$. Its circumcenter is G. Similarly, the centroids G'_m (whose circumcenter is G') for the triangles $H_iH_jH_k$ determine the other quadrisection point S' of OH. Furthermore, N is the center of symmetry for the two quadrilaterals $N_1N_2N_3N_4$ and $N'_1N'_2N'_3N'_4$ and also for $G_1G_2G_3G_4$ and $G'_1G'_2G'_3G'_4$.

There are eight *orthocentroidal circles* (see Droussent) on the segments G_iH_i and on $G'_iH'_i$ as diameters, pairs of which determine 16 distinct radical axes all passing through N, so N is the center of a circle orthogonal to all these eight circles.

We see that the Euler segment could well be renamed the *seven point line* (points O, S, G, N, G', S', H). With this notation, since points G and N trisect and bisect OH, the resemblance to the Euler line of a triangle is striking.

See also H. G. Forder, *Higher Course Geometry*, Cambridge University Press, 1949, 232–235, and R. A. Johnson, *Modern Geometry*, Houghton-Mifflin, 1929, pp. 169, 207, 243, and 251–253.

An Application of Ceva's Theorem

E 2312 [1971, 793]. *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey*

Let D be a point in the plane of a positively oriented triangle ABC and let AD, BD, CD intersect the respective opposite sides in A_1, B_1, C_1. If the oriented segments $\overline{BA_1}$, $\overline{CB_1}$, $\overline{AC_1}$ are equal ($=\delta$), then D is uniquely determined and lies in the interior of ABC. (Notice the analogy between D and the Brocard point Ω.)

Solution by Michael Goldberg, Washington, D.C. Let the lengths of the sides of the triangle be a, b, c, where $a \leqq b \leqq c$. Then by Ceva's Theorem, we have the equation

$$(*) \qquad (a-\delta)(b-\delta)(c-\delta) = \delta^3.$$

The left member of (*) is a function which decreases monotonically from abc at $\delta = 0$ to zero at $\delta = a$, and the right member is a function which increases monotonically from zero at $\delta = 0$. Hence the two functions are equal for exactly one real value of δ which lies in the interval $(0, a)$; it is easy to see also that there are no other real solutions to (*).

Note that if, instead, the segments CA_1, BC_1, and AB_1 are equal, then the value of δ is the same, but the transversals cross at another point E. The points D and E coincide only for the equilateral triangle.

Also solved by Bernhard Andersen (Denmark), Harold Donnelly, Jordi Dou (Spain), M.G. Greening (Australia), V. F. Ivanoff, and the proposer.

Editor's Comment. L. Goldstone located a complete discussion of this point, its isotomic conjugate, and their properties in Peter Yff, *An analogue of the Brocard Points*, this MONTHLY 70 (1963), 495–501.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before December 31, 1972. *Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.*

An asterisk (*) *means neither the proposer nor the editors supplied a solution.*

Editorial. More good proposals for this Section are needed; the supply has been embarrassingly low in recent months. Many proposals are received; but, in the opinions of the Editors, few are acceptable. Too many are either routine exercises or consist of unmotivated strings of definitions and axioms in abstract settings; untangle the terminology and you have solved the problem. A "good" problem, in the subjective opinions of the present Editors, should be brief in statement, understandable, and intriguing — even surprising — to every mathematician. Its solution should depend on at least one ingenious idea or trick, and should be elegant and brief. Textbook exercises that yield to direct attack are unsatisfactory, as are problems comprehensible only to a few specialists.

5866. *Proposed by Hal Forsey, San Francisco State College*

Let A be a nonempty proper subset of R, the real numbers. Show that $\{A + t : t \text{ in } R\}$ is infinite.

5867*. *Proposed by D. E. Daykin, Reading University, England*

Let Q be the real quadratic form

$$\sum_{i=1}^{4}\sum_{j=1}^{4} a_{ij}x_ix_j \qquad \text{with } a_{ij}=a_{ji}.$$

How can we ensure that $Q \geqq 0$ whenever all $x_i \geqq 0$?

5868. *Proposed by B. C. Anderson, Henry Ford Community College*

Show that the following theorem becomes false if "Archimedean" is omitted. R^n is an Archimedean vector lattice with respect to the order generated by a cone K if and only if there are n linearly independent vectors $v^{(k)}$ such that $K=\{x=(x_j)\in R^n\colon \sum_{j=1}^n x_jv_j^{(k)}\geqq 0;\ k=1,2,\cdots,n\}$. (Note: the word "Archimedean" is inadvertently omitted on p. 9 of A. L. Peressini, *Ordered Topological Vector Spaces.*)

5869. *Proposed by Anatol Rapoport, University of Toronto*

Let n points be chosen at random on the circumference of a unit circle. Show that the expected area of the inscribed n-gon is given by

$$A(n)=\pi\left[1-n!\sum_{j=1}^{\infty}\frac{(2\pi)^{2j}(-1)^{j-1}}{(n+2j)!}\right].$$

5870. *Proposed by D. J. Lutzer and F. G. Slaughter, Jr., University of Pittsburgh*

For which discrete spaces D is βD hereditarily normal? (βD denotes the Stone-Čech compactification of D.)

5871*. *Proposed by P. R. Chernoff, University of California, Berkeley*

Let $f(x,y)$ be a real-valued function of two real variables which is separately differentiable. Assume that $\partial f/\partial x=\partial f/\partial y$ everywhere. Must there be a function g of one variable such that $f(x,y)=g(x+y)$? What if we also assume *a priori* that f is jointly continuous?

SOLUTIONS OF ADVANCED PROBLEMS

Universal Covering Series

5763 [1970, 1015]. *Proposed by J. T. Rosenbaum, University of Pittsburgh*

For any set S of reals, call the convergent series $\sum a_n$ an S series if for each $\varepsilon>0$ there exists a sequence $\{I_n\}$ of open intervals covering S with $|I_n|\leqq\varepsilon a_n$, $n=1,2,\cdots$. Find a Cantor set series (see Problem 5665 [1970, 411]). Is there an S series for all S of measure 0? Is there a universal series?

I. *Solution by Nicholas Passell, University of Florida.* In the usual construction of the Cantor set the residual set has at the nth stage 2^n components of length $1/3^n$. Hence we shall succeed in covering the set if for some $k_\varepsilon \geqq 2^n$, $\varepsilon \cdot a_{k_\varepsilon} > 1/3^n$. That is, $\varepsilon a_{k_\varepsilon} > k^{-\log_2 3}$. Take $0 < \delta < (\log_2 3 - 1)$, then $\sum k^{-(\log_2 3 - \delta)}$ is a Cantor set series because k^δ is eventually greater than ε.

II. *Solution by E. Boardman, Westfield College, London, England.* We shall show that if $\sum_{n=1}^\infty a_n$ is a convergent series of positive numbers then there exists a set S of measure zero for which $\sum a_n$ is not an S series. First construct a continuous monotonic increasing function g on the non-negative real line with $g(t) > 0$ for $t > 0$ and $g(0+) = g(0) = 0$ such that

$$(1)\ \lim_{t\to 0} \frac{t}{g(t)} = 0, \qquad (2)\ A = \sum_{n=1}^\infty g(a_n) < \infty.$$

By Dvoretsky's result, (Proc. Camb. Phil. Soc., 44, 1948), one can construct a set S on the real line with $A < g^*m(S) < \infty$. Then (1) implies that $m(S) = 0$. (Here $g^*m(S) = \lim_{\delta\to 0+} g^{\delta *} m(S)$, $g^{\delta} * m(S) = \inf\{\sum_{n=1}^\infty g(d(I_n)) : \bigcup_{n=1}^\infty I_n \supset S, d(I_n) < \delta\}, I_n$ are open intervals and $d(I)$ denotes the diameter of I.)

Let $\varepsilon > 0$ be such that $A < g^*m(S) - \varepsilon$. For small δ, $g^* m(S) > g^* m(S) - \varepsilon > A$, so that if $\bigcup_{n=1}^\infty I_n \supset S$ and $|I_n| < \delta$ for all n, then

$$\sum_{n=1}^\infty g(d(I_n)) > A = \sum_{n=1}^\infty g(a_n).$$

Hence there exists n such that $g(|I_n|) > g(a_n)$ which, as g is monotonic increasing, implies

$$(3) \qquad |I_n| \equiv d(I_n) > a_n.$$

Let $a = \max\{a_n : n = 1, 2, \cdots\}$ and let α be such that $1 \geqq \alpha > 0$ and $\alpha a < \delta$. Then it follows easily that if $\bigcup_{n=1}^\infty I_n \supset S$ there exists n such that $|I_n| > \alpha a_n$, for otherwise one gets a contradiction to (3). So $\sum a_n$ is not an S series.

Deformation Retracts

5765 [1970, 1115]. *Proposed by Simeon Reich, Israel Institute of Technology, Haifa*

On p. 325 of Dugundji, *Topology*, the following result is stated: A is a deformation retract of B over X if and only if A is a retract of B and B is deformable into A over X.

Is this true?

Solution by F. Cunningham, Jr., Bryn Mawr College. No. Let C be a circle

with c a point of C, and let I be the unit interval. Let $X = C \times I$, and let $C_0 = C \times \{0\}$, $C_1 = C \times \{1\}$, be the lower and upper edges of X. Let $B = C_0 \cup C_1$ and let $A = \{(c,0)\} \cup C_1$, so that $A \subset B \subset X$. Then A is a retract of B, for define $r: B \to A$ to be the identity on C_1 and to smash C_0 into the point $(c,0)$ of A. Also B is deformable into A over X by the map $\Phi: B \times I \to X$ such that $\Phi(b,t) = b$ for $b \in C_1$, $t \in I$, and $\Phi(b,t) = (x,t)$ for $b = (x,0) \in C_0$, $t \in I$. But A is not a deformation retract of B over X. Indeed, if $\Psi: B \times I \to X$ is a homotopy from the identity of B to a retraction of B on A, then the restriction of Ψ to $C_0 \times I$ is a homotopy from the identity of C_0 to a loop in A based at $(c,0)$. But the only loop in A based at $(c,0)$ is the trivial one, which is not homotopic in X to the identity of C_0.

Also solved by Jan Hejeman (Czechoslovakia), and by John Henze.

Universal Computer for Continuous Functions

5770 [1970, 1116]. *Proposed by J. P. Jones, University of Calgary*

Does there exist a function $\psi = \psi(x,y)$ continuous on some domain in $\mathbb{R}^2$ with the following property? For each real valued $\theta(x)$ continuous on a connected domain, there exists a y_0 such that $\psi(x,y_0)$ and $\theta(x)$ have the same domain and agree there.

Solution by P. R. Chernoff, University of California, Berkeley. We shall construct such a function ψ with domain a certain G_δ subset of the plane. We shall need the fact that if $\mathcal{N}$ is the space of irrational numbers between 0 and 1 then there is a continuous map from $\mathcal{N}$ onto any Polish space (homeomorph of a complete separable metric space). For a proof of this see, e.g., Parthasarathy, *Probability Measures on Metric Spaces*, Chapter 1.

Let $C(I)$ be the space of continuous real functions on the open unit interval I, with the topology of uniform convergence on compact sets. This is a Polish space, so there is a continuous surjection $T: \mathcal{N} \to C(I)$.

Let $P \subset \mathbb{R}^2$ be the set of all pairs (a,b) with $a < b$, and let U be the open subset of $\mathbb{R} \times P \times \mathcal{N}$ consisting of all $(x;(a,b),n)$ such that $a < x < b$. Define $F: U \to \mathbb{R}$ by

$$F(x;(a,b),n) = T(n)\left(\frac{x-a}{b-a}\right).$$

The continuity of T readily implies that F is continuous on U. Moreover, it is clear that given $a < b$ and a continuous function $\theta:(a,b) \to \mathbb{R}$, there is $n \in \mathcal{N}$ such that

$$\theta(\cdot) = F(\cdot\,;(a,b),n).$$

Now $P \times \mathcal{N}$ is Polish, so there is a continuous surjection $\alpha: \mathcal{N} \to P \times \mathcal{N}$. Define $\beta: \mathbb{R} \times \mathcal{N} \to \mathbb{R} \times P \times \mathcal{N}$ by $\beta(t,n) = (t;\alpha(n))$. Let $E = \beta^{-1}(U)$. E is open in $\mathbb{R} \times \mathcal{N}$ and therefore is a G_δ subset of the plane.

We define ψ on E by

$$\psi(x,n) = F(\beta(x,n)) = F(x;\alpha(n)).$$

Then ψ is continuous and obviously a suitable horizontal section of ψ represents any given continuous real function θ whose domain is a finite open interval.

The domain of ψ is contained in a horizontal strip of height 1. Eight similar constructions in disjoint strips give us continuous functions whose sections represent all continuous real functions on each of the eight other types of connected interval. We piece all nine of them together for our final function ψ.

Finite Cyclic Groups

5774 [1971, 84; 1972, 191]. *Proposed by J. C. Owings, Jr., University of Maryland*

Let G be a finite group and suppose, for all $d \geqq 1$, that G has at most d elements of order d. Prove G is cyclic.

II. *Solution by John Woods, Student, Florida State University.* Let p be any prime divisor of the order of G and S any corresponding Sylow subgroup (with order p^n). By hypothesis, S contains at most p^m elements of order p^m, $0 \leqq m \leqq n$, and since $1 + p + p^2 + \cdots + p^{n-1} < p^n$, S contains at least one element of order p^n. Hence S is cyclic and contains $p^n - p^{n-1}$ elements of order p^n. By Sylow's theorem G contains $1 + kp$ subgroups of order p^n with distinct generators if $k > 0$, and thus $(1 + kp)(p^n - p^{n-1})$ elements of order p^n. But for $k > 0$, $(1 + kp)(p^n - p^{n-1}) > p^n$. Thus $k = 0$ and S is unique. Hence S is normal. Since the Sylow subgroup is normal and cyclic, G is a direct product of its Sylow subgroups and is cyclic.

Editorial Note. Both the proposer and D. M. Bloom have pointed out that the solution which appeared [1972, 191] and also the references cited in Scott, Fraleigh, and Cohn (a) were for a weaker theorem, viz.:

Let G be a finite group and suppose, for all $d \geqq 1$, that G has at most d elements of order dividing d. Then G is cyclic.

Bloom has proved that the weaker theorem also implies the original problem.

Irreducible Polynomials

5780 [1971, 203]. *Proposed by W. R. Emerson, New York University*

For which algebraic number fields $F([F:Q] < \infty)$ is the following valid? A primitive polynomial $P \in \theta[x]$ is reducible over $F[x]$ if and only if it is reducible over $\theta[x]$, where θ is the ring of integers of F.

I. *Solution by W. C. Waterhouse, Cornell University.* The statement is valid if and only if θ is a principal ideal domain. Indeed, if θ is a principal ideal domain, it has unique factorization, and the statement is true by Gauss's lemma. Conversely, suppose I is a nonprincipal ideal; as θ is a Dedekind domain, we can find two elements a, b in θ generating I. Let $I^n = c\theta$ be the first power of I which is principal; such an n exists because the ideal class group is finite. Then $P(x) = (1/c)(ax+b)^n$ is in $\theta[x]$ and is primitive, i.e., its coefficients generate all of θ. If $P(x)$ is reducible in $\theta[x]$, its factors must also be primitive. But any nontrivial factor has the form $d(ax+b)^m$ with $1 \leqq m < n$; the coefficients of this generate the ideal dI^m, which cannot equal θ since I^m is not principal.

II. *Solution by Robert Gilmer, Florida State University.* It is possible to prove the following:

Let F be an algebraic extension field of Q (possibly infinite dimensional), and let θ be the integral closure of Z in F.

(a) *If each finitely generated ideal of θ is principal* (that is, θ is a Bezout domain; for $[F:Q] < \infty$, this means that θ is a principal ideal domain), *then each polynomial $f(X) \in \theta[X]$ irreducible over θ is irreducible over F.*

(b) *If some finitely generated ideal of θ is not principal, then there is a quadratic polynomial $g(X) = aX^2 + bX + c$ in $\theta[X]$ such that $(a,b,c) = \theta$, $g(X)$ is irreducible in $\theta[X]$, and $g(X)$ is reducible in $F[X]$.*

Proof of the above is effected using the two theorems:

(1) *Let R be a Prüfer domain with quotient field S. If R is a Bezout domain, and if n is a positive integer, then each element f of $R[X_1, \cdots, X_n]$ irreducible in $R[X_1, \cdots, X_n]$ is also irreducible in $S[X_1, \cdots, X_n]$. If R is not a Bezout domain, then there is a quadratic polynomial $g(X_1) = aX_1^2 + bX_1 + c$ in $R[X_1]$ such that $(a,b,c) = R$, g is irreducible in $R[X_1]$, and g is reducible in $S[X_1]$.*

(2) *Let R be a Prüfer domain with quotient field K. Let L be an algebraic extension field of K, and let T be the integral closure of R in L. Then T is a Prüfer domain. (In particular, θ is a Prüfer domain.)*

(1) follows from the proof of Theorem 3 of J. Arnold and R. Gilmer, *On the contents of polynomials*, Proc. A.M.S., 24 (1970), 556–562. (2) appears as Theorem 101 of I. Kaplansky, *Commutative Rings*, Allyn and Bacon, 1970.

Also solved by J. W. Brewer & William Heinzer, Mark Yu, and the proposer.

Editorial Note. Gilmer notes that the stated result requires the polynomial P to be primitive. ($2X$ is irreducible in $F[X]$ but is reducible in $\theta[X]$.) The following result is implicit in the solutions received:

Let θ be a Dedekind domain and let F be its quotient field. Then each irreducible polynomial in $\theta[X]$ is irreducible in $F[X]$ if and only if θ is a unique factorization domain (or, equivalently, a principal ideal domain).

A Subquasigroup Generated from $(xy)(y(z(zx))) = y$

5793 [1971,411]. *Proposed by N. S. Mendelsohn, University of Manitoba.*

Let G be a quasigroup with at least two elements and let G satisfy the law

$$(xy)(y(z(zx))) = y \tag{1}$$

for all x, y, z in G. Show that any two distinct elements of G generate a subquasigroup of order 5.

Solution by D. A. Leonard, Ohio State University. In (1) replace x by zx and y by xz [in symbols: $x \to zx$, $y \to xz$] to get

$$[(zx)(xz)][xz(z(z(zx)))] = xz,$$

whence using (1) and $y \to z$, we have

$$[(zx)(xz)]z = xz. \tag{2}$$

Using cancellation in the quasigroup G, we get

$$(zx)(xz) = x. \tag{3}$$

Now (2) with $z \to x$ and cancellation gives

$$(x^2 \cdot x^2)x = x^2, \tag{4}$$

$$x^2 \cdot x^2 = x. \tag{5}$$

Using (1) with $y \to x^2$, $z \to x$, we have $[x(x^2)]\ [x^2(x(x(x^2)))] = x^2$ and (1) with $y \to x$, $z \to x$ reduces this to $[x(x^2)]x = x^2$ or

$$x(x^2) = x \tag{6}$$

and from (5) and (6),

$$x = x^2. \tag{7}$$

Thus G is idempotent. It can easily be seen from (7) and cancellation that if x and y are distinct elements of G, then x, y, xy, yx, and $(xy)x$ are distinct, so it suffices to show closure by completing the multiplication table.

(1), with $y \to x$, $z \to y$, implies $x^2(x(y(yx))) = x$. From (5) and cancellation we have $x(y(yx)) = x^2$ and

$$y(yx) = x. \tag{8}$$

(3) and (8) give $[(xy)x]\ [x(xy)] = x$ and

$$[(xy)x]y = x. \tag{9}$$

Similarly $[x(xy)]\,[(xy)x] = xy$ and

$$(10) \qquad y[(xy)x] = xy.$$

Now (1), with $x \to yx$, $y \to x$, $z \to xy$, gives

$$[(yx)x]\,[x(xy((xy)(yx)))] = x$$

and $[(yx)x]\,[x((xy)y)] = x$ and $x[(xy)y] = x(yx)$, whence

$$(11) \qquad (xy)y = yx.$$

From (9) with $x \to xy$, $y \to x$, we have $(((xy)x)xy)x = xy$ and

$$(12) \qquad ((xy)x)xy = yx.$$

With $y \to xy$, (12) gives $[(x(xy))x]\,[x(xy)] = (xy)x$, whence

$$(13) \qquad (yx)y = (xy)x.$$

Further

$$(14) \qquad (yx)\cdot[(xy)x] = yx[(yx)y] = y,$$

$$(15) \qquad (xy)x \cdot yx = xy,$$

$$(16) \qquad y(xy) = (xy)x,$$

$$(17) \qquad x(yx) = (yx)y = (xy)x,$$

$$(18) \qquad x[(xy)x] = [x(xy)]x = yx.$$

Basically (9) through (18) are reletterings of (3) and (8) with cancellation. These eighteen are sufficient to construct the multiplication table, with which the proof is complete.

$\cdot$	x	y	xy	yx	$(xy)x$
x	x	xy	y	$(xy)x$	yx
y	yx	y	$(xy)x$	x	xy
xy	$(xy)x$	yx	xy	y	x
yx	xy	$(xy)x$	x	yx	y
$(xy)x$	y	x	yx	xy	$(xy)x$

Also solved by M. G. Greening (Australia), R. Padmanabhan (as a corollary in a paper, *Characterization of a class of groupoids*, submitted for publication in *Algebra Universalis*), by L. E. Shader, and by the proposer.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 NUMBER 8

CONTENTS

(*Continued on inside cover*)

OCTOBER 1972

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

THE HISTORICAL DEVELOPMENT OF ALGEBRAIC GEOMETRY

J. DIEUDONNÉ, University of Nice, France and University of Maryland

I. THEMES AND PERIODS

Modern algebraic geometry has deservedly been considered for a long time as an exceedingly complex part of mathematics, drawing practically on every other part to build up its concepts and methods and increasingly becoming an indispensable tool in many seemingly remote theories. It shares with number theory the distinction of having one of the longest and most intricate histories among all branches of our science, of having always attracted the efforts of the best mathematicians in each generation, and of still being one of the most active areas of research. Both are perhaps the best candidates for the perfect mathematical theory, according to Hilbert's ideas: if we agree with him that problems are the lifeblood of mathematics, then certainly we may say that algebraic geometry and number theory always have had more open problems than solved ones, and that each progress towards their solution has always brought with it a host of new and exciting methods.

Human minds being unable to grasp complex matters as a whole, I have thought it would be helpful to describe the history of algebraic geometry as a kind of two-dimensional pattern, where many varied trends of thought, belonging to a few big *themes*, weave their way as multicolored threads through the moving succession of years. It should, however, be emphasized from the start that such a presentation inevitably inflicts distortions on reality: these themes constantly react on one another, and any division of time into periods is bound to founder on the fact that periods almost always overlap.

With these reservations, we may first group the main ideas of algebraic geometry as follows:

(A) and (B) The twin themes of *classification* and *transformation*, hardly to be separated, since the general idea behind classification of algebraic varieties is to put together those which can be deduced from each other by some kind of "trans-

---

It is hardly necessary to identify Prof. Dieudonné to our readers; still a few facts may prove interesting. Prof. Dieudonné studied at the Ecole Normale supérieure from 1924–27, was a Fellow at Princeton, Berlin, and Zürich, and received his Doctorate in 1931. He served on the faculties at Bordeaux, Rennes, Nancy, São-Paulo, Michigan, Northwestern, l'Institut des Hautes Etudes Scientifiques, and was the Dean of Faculty at Nice until his retirement. He held Visiting Professorships at Columbia, Johns Hopkins, Rio de Janeiro, Buenos Aires, Pisa, Maryland, Tata Institute Bombay, Notre Dame, and Washington. His honors include the Order of the Legion of Honor, the Order of the Academic Palms, and membership in the Academy of Sciences. He served as President of the Mathematical Society of France in 1964–65.

Prof. Dieudonné has published a number of books and about 135 research articles on analysis, topology, spectral theory, classical groups, formal Lie groups, and non-commutative rings. This article was prepared while the author was a Visiting Professor at the University of Maryland. *Editor.*

formation." Subordinate to these themes are the notion of *invariant*, both of algebraic type and of numerical type (such as dimension, begree, genus, etc.), and the concepts of *correspondence* and of *morphism*, which give precise meanings and extensions to the vague idea of "transformation."

(C) *Infinitely near points*: a thorny problem, which has plagued generations of mathematicians: the definition and classification of singularities, the correct definition of "multiplicity" of intersections, later the concept of "base points" of linear systems, and the recent introduction of rings with nilpotent elements, all belong to that theme.

(D) *Extending the scalars*: a giant step forward in the search for simplicity: the introduction of *complex* points and later of *generic* points were the forerunners of what we now consider as perhaps the most characteristic feature of algebraic geometry, the general idea of *change of basis*.

(E) *Extending the space*: another fruitful method for extracting understandable results from the bewildering chaos of particular cases: *projective geometry* and *n-dimensional geometry* paved the way for the modern concepts of "abstract" varieties and *schemes*.

(F) *Analysis and topology in algebraic geometry*. This theme beautifully exemplifies the cross-fertilization between various branches of mathematics. Out of a problem of integral calculus, the computation of elliptic integrals and of their generalizations, adelian integrals, Riemann developed the concept of Riemann surface (the first non-trivial example of "complex manifold"), invented algebraic topology, and he and his successors showed how these ideas completely renewed the theory of algebraic curves and surfaces. One hundred years later, history repeated itself when A. Weil transferred to algebraic geometry the notion of *fiber bundle*, and Serre the idea of using *sheaves* and their cohomology, which he and H. Cartan had shown to be so effective for complex manifolds.

(G) *Commutative algebra and algebraic geometry*. As we shall see, this has grown into the most important theme for modern algebraic geometry. Since Riemann introduced the field of rational functions on a curve, Kronecker, Dedekind and Weber the concepts of ideals and divisors, commutative algebra has become the workshop where the algebraic geometer goes for his main tools: local rings, valuations, normalization, field theory, and the most recent and most efficient of all, homological algebra.

Needless to say, within the scope of this article, it will be impossible to do more than deal with a few of the highlights of our history, leaving aside a large number of important developments which should be included in a reasonably complete survey.

II. FIRST PERIOD: "PREHISTORY" (CA. 400 B.C.–1630 A. D.)

If it is true that the Greeks invented geometry as a deductive science, they never (contrary to popular beliefs) made any attempt to divorce it from algebra. On the contrary, one of their main trends was to use geometry to solve algebraic problems, and this is best exemplified in the invention of the conics, the first curves which they thoroughly studied after straight lines and circles. The Greeks knew simple geometric constructions for the root of the equation $x^2 = ab$, a and b being given as *lengths* of segments, and the unknown x being considered as the side of a square; they usually wrote the equation as a "proportion" $a/x = x/b$. The "Delic problem" called for construction of a length x of given cube, $x^3 = a^2b$; this was transformed by Hippocrates of Chio (around 420 B.C.) into a "double proportion" $a/x = x/y = y/b$ for two unknown lengths x, y. Menechmus (ca. 350 B.C.) had the idea of considering the loci given by the two equations $ay = x^2$ and $xy = ab$, whose intersection has as coordinates x, y a solution of the problem. This may seem to involve knowledge of analytic geometry; actually the Greeks made extensive use of coordinates (in particular for the later theory of conics by Apollonius), without however reaching the general point of view of Descartes and Fermat (see below).

This method of solving equations by intersections of curves had in fact already been used in the 5th century B.C., and led to the invention of many curves, both algebraic and transcendental; of course, the distinction between the two kinds of curves could not be perceived during that period, and more generally, there was no attempt at classification, for which no rational basis existed. Besides planes and spheres, the Greeks also studied some surfaces of revolution, such as cones, cylinders, a few types of quadrics and even tori; after having discovered conics "analytically," Menechmus was also the first to recognize that they could be obtained as plane sections of a cone of revolution; and a bold construction of Archytas (late 5th century B.C.) gave a solution of the Delic problem by the intersection of a cone, a cylinder and a torus. Finally, in his astronomical work, Eudoxus was led to describe the intersection of a sphere and a cylinder as the trajectory of a movement conceived as the superposition of two rotations, which may be considered as the first example of a parametric representation of a curve.

III. SECOND PERIOD: "EXPLORATION" (1630–1795)

For once, this period has a very well-defined starting point, the independent invention by Fermat and Descartes of "analytic geometry," which certainly also marks the true birth of algebraic geometry. The main novelty compared to the way the Greeks used coordinates is that the *same* axes are used for *all* curves (fixed

or variable) which are being considered in a problem, and above all the fact that the algebraic notation of Viète and Descartes opens the way to the consideration of arbitrary equations (where the Greeks could not go beyond the third or fourth degree). Within this frame, the distinction between algebraic and transcendental curves immediately emerges; the concept of dimension is already clear to Fermat, who explicitly states that a single equation defines a curve in 2 dimensions, a surface in 3 dimensions, and already hints at the possibility of generalization to higher dimensions. The degree of a plane curve is at once seen to be invariant with respect to a change of coordinates, and Newton knows that it is also invariant under a central projection (an operation which was familiar since the study of conic sections by the Greeks).

Themes A and B

The chief work of that period is one of exploration. Fermat shows that all curves of degree 2 are conics, and Newton classifies all plane cubics with respect to change of coordinates and projections; Euler classifies the quadrics, and the first skew curves, given as intersection of two surfaces, appear in the 18th century. The concept of parametric representation of a curve is fundamental in Newton's approach to calculus, and Euler knows how to get in certain cases a parametric representation from the cartesian equation. A beginning is made in the elucidation of the structure of singular points and inflexion points of algebraic plane curves, although limited to the most elementary cases, so that no general description is yet obtained.

Theme C

The problem of intersection of two algebraic plane curves is already tackled by Newton; he and Leibniz had a clear idea of "elimination" processes expressing the fact that two algebraic equations in one variable have a common root, and using such a process, Newton observed that the abscissas (for instance) of the intersection points of two curves of respective degrees m, n, are given by an equation of degree $\leqq mn$. This result was gradually improved during the 18th century, until Bézout, using a refined elimination process, was able to prove that, in general, the equation giving the intersections had exactly the degree mn; however, no general attempt was yet made during that period to attach to each intersection point an integer measuring the "multiplicity" of the intersection, in such a way that the sum of the multiplicities should always be mn. Bézout also generalized his elimination process to 3 dimensions, proving that the points of intersection of three algebraic surfaces of degrees m, n, p are in general given by an equation of degree mnp.

With the beginning of the consideration of algebraic families of algebraic curves a problem in a sense converse to the problem of intersections appeared, namely the determination of a curve of given degree n containing sufficiently many given points. It should be recalled here that this (linear) problem was the starting point for the theory of determinants, and the fact that $n(n+3)/2$ points in "general position"

completely determine a curve of degree n, whereas two curves of degree n have in general n^2 common points, gave the first general example of the concept of rank for a system of linear equations ("Cramer's paradox").

We should finally stress the fact that a number of ideas fully developed during the next period may be traced back (in an embryonic form) to the 17th or 18th century, as we shall see below.

IV. THIRD PERIOD: "THE GOLDEN AGE OF PROJECTIVE GEOMETRY" (1795–1850)

Here again we have a rather sharp break with the past at the beginning of this period. In the space of a few years, with Monge and his school and especially with Poncelet, a new era begins with the simultaneous introduction of points at infinity and of imaginary points: "geometry" will now, for almost 100 years, exclusively mean geometry in the complex projective plane $\boldsymbol{P}_2(\boldsymbol{C})$ or the complex projective 3-dimensional space $\boldsymbol{P}_3(\boldsymbol{C})$. In fact, the fundamental idea of (real) projective geometry goes back to Desargues (17th century) who, trying to give mathematical foundations to the methods of "perspective" used by painters and architects, had introduced the concept of "point at infinity," and the use of central projections as a means of getting new theorems from classical results of Euclidean geometry; and although these ideas had inspired Pascal in his work on conics, they had very soon dropped into oblivion, due to the outlandish language of the author and the very limited diffusion of his book (which was for some time believed lost). Other mathematicians in the 18th century, in particular Euler and Stirling, had hinted at the existence of imaginary points, in order to state general theorems without distinction of various cases. This is precisely what is brilliantly accomplished by the new school: circles now intersect in 4 points as any two conics should, but two of the points are imaginary and at infinity; instead of several kinds of conics and quadrics, all nondegenerate conics (resp. quadrics) are now projectively equivalent; instead of the 72 kinds of cubics enumerated by Newton, only 3 remain projectively distinct; etc.

Themes D and E

The chief beneficiaries of these new ideas are at first the theory of conics, quadrics and of linear families of conics and quadrics; but curves and surfaces of degree 3 or 4 are also investigated in this way, revealing beautiful new theorems, such as the configurations of the 9 inflexion points of a plane cubic, the 27 lines on a cubic surface, the 28 bitangents to a plane quartic; the theorem of Salmon, proving the constancy of the cross ratio of the 4 tangents to a cubic issued from a point of the curve, was to gain even more significance later, as the first concrete example of a "module" in Riemann's sense for an algebraic curve.

Although, with Möbius, Plücker and Cayley, projective geometry received a sound algebraic basis by the use of homogeneous coordinates, a general tendency

of the projective school was to minimize as much as possible algebraic computations, and to rely instead (beginning with Poncelet) on general heuristic "principles" which they did not bother to justify algebraically. The remarkable success they had in this direction was chiefly due to their skillful use of the idea of geometric *transformation*, which for the first time comes to the forefront in geometry, preparing the ground for Klein's famous "Program" linking geometry and the theory of groups. Most of the transformations they consider are linear: for instance, one of their favorite devices in the theory of conics is to consider a conic as the locus of two variable straight lines through two fixed points, one of them being derived from the other by a fixed linear transformation (an idea which, in some particular cases, goes back to Maclaurin). Similarly, in the study of the linear system of conics through 4 fixed points, they investigate the intersections of these conics with a fixed straight line D by considering the (linear) transformation which to a point M of D associates the second point of intersection with D of the conic of the system which contains M. Emboldened by the results obtained in this manner, they inaugurated what was to become the theory of *correspondences*, by considering what they called (α,β)-correspondences, i.e., relations between two points M, M' such that to each point M there exist α points M' related to M, and to each point M' there exist β points M related to M': when M and M' vary on the same projective line, Chasles' "correspondence principle" says that the number of points M (counted with multiplicities) coinciding with one of their transforms is $\alpha + \beta$ unless *every* point of the projective line has that property, a result which it is easy to justify algebraically. A beautiful application is the Poncelet "closure theorem" for polygons inscribed in a conic C and circumscribed to a conic C': for a given integer n, one defines on C a $(2,2)$-correspondence by assigning to $M \in C$ the nth point M_n in a sequence $M_0 = M, M_1, \cdots, M_n$, where each side M_iM_{i+1} is tangent to C' and the points M_i are on C. It is easily seen that for n even, one has $M = M_n$ if $M_{n/2}$ is a point common to C and C', and for n odd, $M = M_n$ if $M_{(n-1)/2} = M_{(n+1)/2}$, and the tangent to C at that point is also tangent to C'. There are thus at least 4 points M on C such that $M = M_n$, and by the correspondence principle, if there is still *one* more point having that property, then $M = M_n$ for *all* points on C (one uses of course the parametrization of a conic by the projective line).

Theme B

Later representatives of the projective school (notably Chasles in France, Steiner and von Staudt in Germany), somewhat intoxicated by the elegance of their methods, went so far as to insist that "pure" geometry should be entirely divorced from algebra and even (with von Staudt) from the concept of real number. As could be expected, such efforts did not lead very far, and probably hampered progress in the realization of the importance of linear algebra in classical geometry; it may be, however, that they paved the way for the later "abstract" algebraic geometry over a field different from $\boldsymbol{R}$ or $\boldsymbol{C}$.

In the general theory of algebraic curves (in $\boldsymbol{P}_2(\boldsymbol{C})$) and surfaces (in $\boldsymbol{P}_3(\boldsymbol{C})$), the main problems studied before Riemann are of an enumerative character: to give only one example of such problems, what is the number of conics tangent to 5 given conics in general position? (The correct answer is 3264.) Chasles, and later Schubert and Zeuthen proposed half-empirical formulas to solve these problems, based on an intuitive concept of "intersection multiplicity" which could only be justified much later. One of the main ideas of projective geometry, the concept of *duality*, led to the introduction of new "tangential" invariants for algebraic plane curves: the class (number of tangents through a point), the number of inflexion points and the number of double tangents, culminating in the famous "Plücker formulas"

Theme C

Theme A

$$m' = m(m-1) - 2d - 3s,$$

$$m = m'(m'-1) - 2d' - 3s',$$

$$s' - s = 3(m' - m),$$

where m is the degree of the curve, m' its class, d the number of double points, d' the number of double tangents, s the number of cusps, s' the number of inflexion points; no "higher singularities," either punctual or tangential, are supposed to occur.

V. FOURTH PERIOD: "RIEMANN AND BIRATIONAL GEOMETRY" (1850–1866)

The importance of Riemann in the history of algebraic geometry can hardly be overestimated, but in his two fundamental contributions, the "transcendental" approach *via* abelian integrals and the introduction of the field of rational functions on a curve, he built on basic ideas inherited from the previous period.

The origin of abelian integrals is the study of integrals of type

$$\int \frac{R(t)dt}{\sqrt{P(t)}}$$

where $P(t)$ is a polynomial of degree 3 or 4 and $R(t)$ a rational function; one of these integrals expresses the length of an arc of an ellipse (hence the name "elliptic integrals"). In the first half of the 18th century, Fagnano and Euler, looking for some substitute for the classical formula expressing the sum of two arcs of a circle, when the circle is replaced by an ellipse, found indeed that the sum

$$\int_a^x \frac{dt}{\sqrt{P(t)}} + \int_a^y \frac{dt}{\sqrt{P(t)}}$$

can be written

$$\int_a^z \frac{dt}{\sqrt{P(t)}} + V(x, y),$$

where z is an *algebraic* function of x and y, and V a rational or logarithmic function of x and y, and Euler had similar results for more general integrals.

At the beginning of his famous work of elliptic functions, Abel made a giant step forward by showing that the Fagnano-Euler relations were special cases of a very general theorem: he considers an arbitrary "algebraic function" y of x, defined as a solution of a polynomial equation $F(x, y) = 0$; an "abelian integral" $\int R(x, y)dx$ is an integral in which R is a rational function of x, y, in which y is replaced by the preceding algebraic function (for instance elliptic integrals correspond to $F(x, y) = y^2 - P(x)$). Then, if $G(x, y, a_1, \cdots, a_r) = 0$ is a second polynomial in x, y whose coefficients are rational functions of some parameters $a_1, \cdots, a_r$, and if (x_1, y_1), $(x_2, y_2), \cdots, (x_m, y_m)$ are the points of intersection of the two curves $F = 0$, $G = 0$, the sum

$$V = \int_{(a,b)}^{(x_1,y_1)} R(x, y)dx + \cdots + \int_{(a,b)}^{(x_m,y_m)} R(x, y)dx$$

is a rational or logarithmic function of the parameters a_j $(1 \leqq j \leqq r)$*; surprisingly enough, this is little more than an exercise in the theory of symmetric functions of the roots of a polynomial. But Abel does not stop there, and studies in detail the case, in which V is a constant; this leads him to the realization that in that case, *any* sum

$$\int_{(a,b)}^{(x_1,y_1)} R(x, y)dx + \cdots + \int_{(a,b)}^{(x_m,y_m)} R(x, y)dx$$

with *arbitrary* points (x_j, y_j) on the curve $F = 0$, can be expressed as the sum of a *fixed* number δ of values of the same integral, with upper limits algebraic functions of the (x_j, y_j); but, in contrast with the Fagnano-Euler formulas for elliptic integrals, he showed that the number δ may well be > 1, for instance when $F(x, y) = y^2 - P(x)$ with P of degree $\geqq 5$.

Abel, however, worked exclusively within the framework of analysis, and does not seem to have been acquainted with projective geometry. Furthermore, he obviously had no clear concept of integration in the complex plane (in 1826, Cauchy had hardly begun his work on that subject), and with the exception of a short and

---

* Of course, the points x_j, y_j usually have complex coordinates; an integral

$$\int_{(a,b)}^{(x_j,y_j)} R(x, y)dx$$

is only properly defined when the path of integration in the complex plane $\boldsymbol{C}$ with extremities a and x_j has been fixed, and y_j is the value taken by y when x varies along the path, y is a continuous function of x and takes the value b at $x = a$. When the path is replaced by another one (with the same extremities), the value of the integral is modified by a "period."

By definition, a logarithmic function of the a_j has the form log S $(a_1, ..., a_r)$ where S is rational.

inconclusive note, he has no general discussion of the *periods* of his integrals. Thus, although Abel's theorem paved the way for Jacobi's breakthrough in the problem of inversion of hyperelliptic integrals*, Abel himself narrowly missed the concept of integral of the first kind and the definition of the genus of a curve (his failure to take into account the points at infinity has as a consequence the fact that the δ integrals he considers are not necessarily of the first kind).

When Riemann takes up the subject in 1851, the intervening years had seen the great development by Cauchy and his school of the theory of functions of a complex variable. Indeed, the starting point of Riemann has nothing to do with algebraic functions, but is the extension of Cauchy's theory to the "surfaces" he introduces in order better to deal with the so-called "multiform" functions of the most general (not necessarily algebraic) type. This was already far beyond the contemporary concepts, and during the 30 years following Riemann, it was the object of long and tedious explanations by the expositors of his theory. But the way Riemann uses this notion in order to attack the problem of abelian integrals is much more original still. Instead of starting (as would all his predecessors and most of his immediate successors) from an algebraic equation $F(s,z) = 0$ and the Riemann surface of the algebraic function s of z which it defines, his initial object is an n-sheeted Riemann surface without boundary and with a finite number of ramification points**, given *a priori* without any reference to an algebraic equation (Riemann

Theme F

* The natural idea of "inverting" the integral $\int_a^x (Q(t)\,dt)/\sqrt{P(t)} = u$ is to study x as a function of u, as Abel and Jacobi had done when P has degree 3 or 4; but Jacobi realized that, due to the existence of 4 periods, no meromorphic function of u could be a solution of the problem. Abel's theorem finally led him to the correct conception of the problem: one considers *two* equations

$$\int_a^x \frac{dt}{\sqrt{P(t)}} + \int_a^y \frac{dt}{\sqrt{P(t)}} = u, \quad \int_a^x \frac{t\,dt}{\sqrt{P(t)}} + \int_a^y \frac{t\,dt}{\sqrt{P(t)}} = v,$$

and one "inverts" them by expressing the symmetric functions $x + y$ and xy as functions of u and v; Abel's theorem yields an "addition formula" for these functions, from which one can show that they are meromorphic and quadruply periodic.

** The best way to define at least the part of the Riemann surface of a function $s(z)$ (defined by an algebraic relation $F(s, z) = 0$), containing no point at infinity, is to say that it is the subset of C^2 consisting of the pairs (s, z) satisfying the equation $F(s, z) = 0$; there is then no difficulty with the "crossing of sheets." Ramification points are those for which $\partial F/\partial s\,(s, z) = 0$; Puiseux proved in 1850 that if (s_0, z_0) is such a point, the surface decomposes at that point into a finite number of "branches" such that each branch can be represented by equations of type

$$z - z_0 = t^h, \; s - s_0 = a_1 t + a_2 t^2 \ldots,$$

where t (the "uniformizing parameter") is in a neighborhood of 0 in C and the series converges (the integer h depending on the branch).

This description is only correct, however, when at each ramification point (s_0, z_0) there is only one branch; if not, the point (s_0, z_0) must be replaced by as many points as there are branches; in other words the points of a Riemann surface are the *branches* at the various points of the curve.

takes care to complete each sheet with a point at infinity, and thus avoids Abel's difficulties with these points); then he attacks the problem in the most general manner possible: classify the integrals of *all meromorphic functions* on the surface. The work of Cauchy and Puiseux had brought to light the general idea of "periods" of such integrals, generally expressed (as in the example first given by Abel) as an integral taken along an arc joining two ramification points. Here again Riemann breaks entirely new ground: he realizes for the first time that topological concepts are closely related to the problem, and begins by essentially creating the topological study of compact orientable surfaces, attaching to such a surface S an invariantly defined integer $2g$, the minimal number of simple closed curves C_j on S needed to make the complement S' of their union simply connected. Then, instead of studying integrals of meromorphic functions, he *defines directly* integrals of the first and second kinds by their periodicity properties, as functions meromorphic on S', and tending on both sides of each C_j to limits which differ by a quantity k_j constant on C_j (a further reduction of the domain S' is needed to obtain similarly the integrals of the third kind, having logarithmic singularities)*; integrals of the first kind are those which have no pole on S. The existence of integrals of the three kinds is proved by Riemann as a consequence of what he calls the "Dirichlet principle," i.e., the existence of a harmonic function in S' taking prescribed values on the boundary (which allows him to prescribe at will the *real parts* of the k_j); and it is also by an ingenious use of the same principle that Riemann obtains the fundamental relation

$$g - 1 = w/2 - n$$

giving the genus in function of the number of sheets n, and the number w of ramification points (supposed to be of a "general" type).

The meromorphic functions on S are then the integrals of the first or second kind whose periods k_j all vanish, and Riemann shows that they may be expressed as rational functions of two of them, linked by an algebraic relation $F(s, z) = 0$, thus recovering the older point of view, but immeasurably enriched with new insights. The choice of these meromorphic functions s, z is in a large measure arbitrary, and this leads Riemann to his next big step forward, the general concept of *birational transformation* between two irreducible algebraic curves, corresponding to a biholomorphic mapping of their Riemann surfaces. Here again, Riemann was not without predecessors: already Newton and his followers had introduced quadratic transformations such as

Theme B

$$x' = 1/x, \; y' = y/x$$

in the plane, and observed that they thus transformed an algebraic curve into a

* One simply joins the singularity to one of the C_j by an arc, and deletes the arc from S'.

curve of different degree. "Inversion" in the plane and in 3 dimensional space had been intensively studied since the early 1820's, chiefly by "synthetic" geometers; finally, the passage from a plane curve to its transform by duality (exchanging punctual and tangential coordinates) was obviously a birational transformation between two algebraic curves, exchanging degree and class. But the startling novelty of Riemann's approach is of course the fact that to a class of "birationally equivalent" irreducible algebraic curves he was able to attach his topological invariant g, the *genus* of all the curves in the class. But he did not stop there, and by an evaluation (using two different methods) of the parameters on which a Riemann surface of genus g depended, he arrived at the conclusion that classes of isomorphic Riemann surfaces of genus $g \geqq 2$ were characterized by $3g - 3$ complex parameters varying continuously (for $g = 1$ there is only one parameter, and none for $g = 0$); the precise meaning of this result (the so-called theory of "moduli" of curves) was to remain until very recently among the least clarified concepts of the theory.*

Theme A

VI. FIFTH PERIOD: "DEVELOPMENT AND CHAOS" (1866–1920)

The extraordinary wealth of new ideas and methods introduced by Riemann provided inspiration for a steady development of algebraic geometry for over 80 years. But the grandiose synthesis he had envisioned and tried to materialize was almost immediately broken up by his successors. During that period there will be at least two or three schools of algebraic geometry, each using different methods, with little in common even in the fundamental concepts. Riemann's use of analysis, in particular in the "Dirichlet principle," exceeded the possibilities of his time, and he had obviously neglected all the difficulties bound to the existence of singular points on algebraic curves. The first task to which each school of algebraic geometry addressed itself was therefore the systematization of the birational theory of algebraic plane curves, incorporating most of Riemann's results with proofs in conformity with the principles of the school. Then, with varying success, they tried to extend their methods to the theory of algebraic surfaces and higher dimensional algebraic varieties.

VI a: The algebraic approach. Historically, this was the latest one, being initiated by two fundamental papers in 1882, one by Kronecker and one by Dedekind and Weber. But in the light of subsequent history, it is the trend which was to exert the deepest influence on the birth of our modern concepts; in particular, just as Riemann

* One should emphasize the fact that this only describes the first half of Riemann's paper on abelian integrals; the second part, which solves in a masterly way the inversion problem by the introduction of the general "thêta functions" has been, if anything, even more influential on the development of analysis.

had revealed the close relationship between algebraic varieties and the theory of complex manifolds, Kronecker and Dedekind-Weber brought to light for the first time the deep similarities between algebraic geometry and the burgeoning theory of algebraic numbers, which were to be some of the main driving forces during the next periods. Furthermore, this conception of algebraic geometry is for us the clearest and simplest one, due to our familiarity with abstract algebra; but it was precisely this "abstract" character which made it the least popular and least understood one in its time.

The work of Kronecker and of his immediate followers, Lasker and Macaulay, in the first two decades of the 20th century, was of a very general nature, and its importance only emerged in the later periods: it essentially consisted in setting up and consistently using an elimination method, far more flexible and powerful than the preceding ones, with the help of which it was for the first time possible to give a precise meaning to the concepts of *dimension* and of *irreducible variety** and to show that each variety (defined by an arbitrary system of algebraic equations) in projective n-space decomposed in a unique way into a union of irreducible varieties (in general of different dimensions).

Theme G

The goal of Dedekind and Weber in their fundamental paper was quite different and much more limited; namely, they gave purely algebraic proofs for all the algebraic results of Riemann. They start from the fact that, for Riemann, a class of isomorphic Riemann surfaces corresponds to a *field* K of rational functions, which is a finite extension of the field $\boldsymbol{C}(X)$ of rational fractions in one indeterminate over the complex field; what they set out to do, conversely, if a finite extension K of the field $\boldsymbol{C}(X)$ is given *abstractly*, is to reconstruct a Riemann surface S such that K will be isomorphic to the field of rational functions on S. Their very original and fruitful method may be presented in the following way: *if* the Riemann surface S was already known, at each point $z_0 \in S$, a rational function $f \neq 0$ would have an *order* $v_{z_0}(f)$, namely the integer (positive or negative) which is the degree of the smallest power in the Puiseux development $f(u) = \sum_k a_k u^k$ with respect to a "uniformizing parameter" u (equal to $z - z_0$ if z_0 is not a ramification point, to some power $(z - z_0)^{1/h}$ if z_0 is a ramification point). For a *fixed* $z_0 \in S$, the mapping $f \mapsto v_{z0}(f)$ of K^* into $\boldsymbol{Z}$ is what is called a *discrete valuation* on K: we recall that this is by definition a mapping $w\colon K^* \to \boldsymbol{Z}$ such that $w(f+g) \geqq \inf(w(f), w(g))$ if $f+g \neq 0$, and $w(fg) = w(f) + w(g)$, which implies $w(1) = 0$ and $w(f^{-1}) = -w(f)$ (w is usually extended to K by taking $w(0) = +\infty$ by convention). What Dedekind and Weber do is to *reverse* this process, and *define* a "point of the Riemann surface S of K"

* An irreducible variety V in $\mathbf{P}_n(\mathbf{C})$ is characterized by the property that if the product PQ of two homogeneous polynomials is 0 in V, then one of the two polynomials P, Q must be 0 in V. The restrictions to V of the rational functions which are defined at one point of V at least then form a field whose transcendence degree over $\mathbf{C}$ is the dimension of V.

as a *nontrivial discrete valuation* on K (i.e., one which is not identically 0 on K^*: two proportional valuations are then identified).

Now the nontrivial discrete valuations on the field $\boldsymbol{C}(X)$ are easily determined: one of them (the "point at infinity") w_∞ is such that $w_\infty(P) = -\deg(P)$ for any nonzero polynomial $P(X)$; the other ("finite points") correspond bijectively to the points $\zeta \in \boldsymbol{C}$, the corresponding valuation w_ζ being such that $w_\zeta(P)$ is the order of the zero ζ of $P(X)$ (equal to 0 if $P(\zeta) \neq 0$). It can easily be shown that for each discrete valuation w of $\boldsymbol{C}(X)$, there is a finite number of nonproportional valuations v_j *on* K such that for each j, v_j/e_j reduces to w on $\boldsymbol{C}(X)$, where e_j is an integer $\geqq 1$; one says that the v_j are the points of the Riemann surface S *above* w; the points above w_∞ are again called points at infinity, the other finite points.

The elements $f \in K$ for which $v(f) \geqq 0$ for all *finite* points v of S constitute exactly the elements of K which are *integral** over the ring of polynomials $\boldsymbol{C}[X]$; they form what we now call a *Dedekind ring* A, to which Dedekind's theory of *ideals* may be applied.** The maximal ideals $\mathfrak{P}_v$ of A correspond to the finite points $v \in S$: $\mathfrak{P}_v$ is the set of $f \in A$ for which $v(f) > 0$; the *fractionary ideals* of K are the A-modules $\mathfrak{a}$ contained in K and for which there is an element $c \neq 0$ in A such that $c\mathfrak{a} \subset A$; each of them can be written uniquely as a product $\mathfrak{P}_1^{\alpha_1}\mathfrak{P}_2^{\alpha_2}\cdots\mathfrak{P}_r^{\alpha_r}$, where the $\mathfrak{P}_j$ are maximal ideals of A and the α_j positive or negative integers. Another way of stating this result is to say that a fractionary ideal $\mathfrak{a}$ is the set of all $f \in K$ such that $v_j(f) \geqq \alpha_j$ for $1 \leqq j \leqq r$, where the valuations v_j correspond to the maximal ideals $\mathfrak{P}_j$, and $v(f) \geqq 0$ for the other finite valuations.

The consideration of the ideals of A, however, leaves the "points at infinity" out of the picture. This led Dedekind and Weber to generalize the concept of ideal and to introduce the notion of *divisor* on K. This is defined as a family $D = (\alpha_v)$ of integers $\alpha_v \in \boldsymbol{Z}$, where v runs through *all* points of S, and $\alpha_v = 0$ except for a finite number of points: writing $(\alpha_v) + (\beta_v) = (\alpha_v + \beta_v)$ defines the set $\mathscr{D}(K)$ of divisors of K as an *additive group* isomorphic to $\boldsymbol{Z}^{(S)}$, in which an *order relation* is naturally defined, $(\alpha_v) \leqq (\beta_v)$ meaning that $\alpha_v \leqq \beta_v$ for all $v \in S$; a divisor $D = (\alpha_v)$ such that $\alpha_v \geqq 0$ for all $v \in S$ is called *positive* or *effective*. The *degree* $\deg(D)$ of $D = (\alpha_v)$ is defined as $\sum_{v \in S} \alpha_v$ (positive or negative integer); the *support* of D is the set of the $v \in S$ for which $\alpha_v \neq 0$. One of the problems considered by Riemann was the determination of rational functions on a Riemann surface having poles of orders $\leqq \alpha_P$ for prescribed points P (in finite number) on S. Using his bold expression of functions as sums of abelian integrals, he found that there existed rational functions having that property for an *arbitrary* choice of the points P as long as $\sum_P \alpha_P \geqq g + 1$, whereas if $\sum_P \alpha_P \leqq g$, this was only possible for *special* positions of the points P. This result was completed by his student Roch, and put in its final form by Dedekind

* Recall that an element x of a ring R is *integral* over a subring S if it satisfies an equation of type $x^m + a_1 x^{m-1} + \dots + a_m = 0$, with $a_j \in S$.

** Dedekind had developed this theory for algebraic number fields from 1870 on.

and Weber in the following way: the problem is a special case of the study of the set $L(D)$ of rational functions $f \in K$ satisfying the conditions

$$v(f) \geqq -\alpha_v \text{ for all } v \in S \tag{1}$$

for a given divisor $D = (\alpha_v)$; it follows from the axioms of valuations that $L(D)$ is a complex vector subspace of K, and it can be shown that this subspace has *finite* dimension $l(D)$.

A fractionary ideal may be described as the union of the increasing family of spaces $L(D_m)$, where $D_m = (\alpha_v)$ is such that the α_v coincide with the $-\alpha_j$ for the v_j, are equal to 0 for the other finite points, and to m for the points at infinity.

The relations (1) can be written in a different way. For each $f \in K^*$, there are only a finite number of valuations $v \in S$ such that $v(f) \neq 0$; let $(f)_0$ (resp. $(f)_\infty$) be the positive divisor $((v(f))^+)$ (resp. $((v(f))^-)$) (in the "transcendental" interpertation, $(f)_0$ is the "divisor of zeroes" and $(f)_\infty$ the "divisor of poles" of the rational function f), and let $(f) = (f)_0 - (f)_\infty$ in the group $\mathscr{D}(F)$; (f) is called the *principal divisor* defined by f. It can be shown that $\deg((f)) = 0$ by purely algebraic arguments (in the transcendental picture, this is merely the *residue theorem*)*; in particular, if $v(f) \geqq 0$ for *all* $v \in S$, then $f \in C$ (only constants are everywhere holomorphic on a Riemann surface) and if in addition $v(f) > 0$ for some v, then $f = 0$. With these definitions, the relations (1) for $f \neq 0$ are equivalent to the inequality

$$(f) + D \geqq 0 \tag{2}$$

in the ordered group $\mathscr{D}(K)$.

Principal divisors form a subgroup $\mathscr{P}(K)$ of $\mathscr{D}(K)$ (isomorphic to the group K^*/C^*, two elements of K^* which have the same principal divisor differing by a constant factor by the previous remarks). Divisors belonging to the same *class* in the quotient group $\mathscr{C}(K) = \mathscr{D}(K)/\mathscr{P}(K)$ are called (linearly) *equivalent*: to say that D and D' are equivalent means therefore that there exists $f \neq 0$ such that $D' - D = (f)$; it is clear that $\deg(D') = \deg(D)$ and $l(D') = l(D)$ for equivalent divisors; two elements f, g of $L(D)$ are such that $(f) + D = (g) + D$ if and only if f/g is a constant, in other words, the set $|D|$ of *positive* divisors equivalent to D is identified to the projective space $P(L(D))$ of dimension $l(D) - 1$.

The *Riemann-Roch theorem* is then written in the following way:

$$l(D) - l(\Delta - D) = \deg(D) + 1 - g, \tag{3}$$

where g is the genus, and Δ belongs to a well-determined divisor class, called the *canonical class* of K. To define it in the transcendental interpretation, one considers on the Riemann surface S a *meromorphic differential form* ω: at each point P of S,

* One integrates the differential df/f on the boundary of the simply connected part S' of the Riemann surface, taking into account that each arc of that boundary comes twice in the integral with opposite orientations.

the differential form ω may be written $F(u)du$, where u is the uniformizing parameter in a neighborhood of P and F is meromorphic at P; if δ_P is the order of F at the point P, (δ_P) is a *canonical divisor*, and it does not depend on the choice of the uniformizing parameters. Any other meromorphic differential form may be written $f\omega$ with $f \in K$, hence all canonical divisors belong to the same *class*. There is a purely algebraic definition of Δ (see section VII b), and one proves that $\deg(\Delta) = 2g - 2$ for $g \geqq 1$ and $l(\Delta) = g$. Relation (3) implies Riemann's result on the poles of rational functions; more generally, if $\deg(D) \geqq g + 1$, (3) implies $l(D) \geqq 2$; if $D \geqq 0$, $L(D)$ always contains the constant functions, and to say that $l(D) \geqq 2$ means that it contains a non constant rational function. From the definition of $L(D)$, it follows that $l(D) = 0$ if $\deg(D) < 0$, hence, by (3), $l(D) = \deg(D) + 1 - g$ if $\deg(D) > 2g - 2$; in particular, for any divisor D such that $\deg(D) > 0$, $l(mD) = m \cdot \deg(D) + 1 - g$ for m large enough (although one may have $l(D) = 0$).

VI b: The Brill-Noether theory of linear systems of points on a curve. An irreducible plane curve Γ without singularity is identified to its Riemann surface, and a positive divisor may therefore be identified with a system of points of Γ, each being counted with a certain "multiplicity" which is a positive integer. Riemann's determination of the "special" systems of at most g points of Γ, which may be the poles of a rational function, had led him (by an extension of some earlier computations of Abel) to define these sets as intersections with Γ of a family of "adjoint" curves of smaller degree, subject to linear conditions on the coefficients of their equations, so that such a family may be considered as given by an equation $\sum_{j=1}^{r} \lambda_j P_j(x, y) = 0$ in nonhomogeneous coordinates, where the P_j are polynomials and the λ_j variable complex parameters. A number of points of intersection of these curves with Γ may be fixed (i.e., independent of the λ_j); as the intersection multiplicity of a common point of Γ and of an arbitrary curve Γ' is immediately defined since Γ has no singular point (it is the same as the intersection multiplicity of Γ' and the tangent to Γ), we may consider for each adjoint curve Γ' of the family the positive divisor $D = \sum_P m_P P - \sum_Q m_Q^0 Q$, where P runs through *all* the intersection points of Γ and Γ', m_P is the corresponding intersection multiplicity, Q runs through the *fixed* intersection points and m_Q^0 is the minimum value of m_Q when the λ_j vary. It is immediate to see that if D_0 is one of these divisors, corresponding to the values λ_j^0 of the parameters, then $D = D_0 + (f)$, where $f = (\sum_j \lambda_j P_j)/(\sum_j \lambda_j^0 P_j)$.

Conversely, given a divisor D_0 (positive or not), if $l(D_0) = r > 0$, the functions $f \in L(D_0)$ may be written $(\sum_{j=1}^{r} \lambda_j P_j(x, y))/Q(x, y)$, where the P_j and Q are polynomials and the λ_j arbitrary complex numbers; the positive divisors $(f) + D_0$ where $f \in L(D)$, are obtained by adding a fixed divisor to the variable divisor obtained as above from the points of intersection of Γ and of the curve $\sum_j \lambda_j P_j(x, y) = 0$.

The study of the vector spaces $L(D)$ attached to divisors is thus essentially equivalent to the study of the systems of points of intersection (with multiplicities) of Γ with the curves Γ' of a system of curves $\sum_j \lambda_j P_j(x, y) = 0$. It is in fact by means of

the study of such systems of points, called "*linear series*" or "*linear systems*" on Γ, that the geometric school of Clebsch, Gordan, Brill, and Max Noether described the birational theory of algebraic plane curves after 1866. But they wanted to deal in this way, not only with curves without singularities, but with arbitrary algebraic curves, and linear systems of points are only easy to handle when the curve Γ has no singularities, or at most "nice" singularities such as double points with distinct tangents. One of the first efforts of that school was therefore to establish the possibility of finding a birational transformation of an arbitrary irreducible algebraic curve Γ into a plane curve with only double points with distinct tangents; a result proved independently by M. Noether in 1871 and equivalent to a theorem of algebra obtained by Kronecker in 1862. In view of the extension of this result during the later periods, it is worthwhile to note that a slightly weaker theorem may be obtained by a succession of birational transformations of the whole projective plane $\boldsymbol{P}_2(\boldsymbol{C})$ onto itself of the type

Theme C

$$x'/yz = y'/zx = z'/xy$$

(for suitable *homogeneous* coordinates), the so-called *quadratic transformations.* Such a transformation is bijective outside the sides of the triangle having as vertices the points $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ but sends each point of one side (not a vertex) to the opposite vertex, and is indeterminate at a vertex: however, two points approaching a vertex along distinct lines have transforms which tend to distinct limits on the opposite side, so that the transformation may be said to "blow up" a vertex to the opposite side, and *separates* the branches of a curve having different tangents at a vertex by transforming them to branches through different points of the transformed curve. By repeating conveniently this process, one may show that there is a transformed curve whose singular points are such that each has a number of distinct tangents equal to its multiplicity. To get curves with only double points, one uses birational transformations which are only defined on the given curve (and not in the plane).

It is during the same period, and in the same school, that n-dimensional algebraic geometry comes into its own for any value of $n \geqq 1$ (all algebraic varieties being considered as subvarieties of some $\boldsymbol{P}_n(\boldsymbol{C})$). As we shall see below, the study of algebraic varieties of dimension $\geqq 2$ was to have important repercussions on the theory of algebraic curves, with the concept of algebraic correspondences as subvarieties of a product variety, and the study of abelian varieties. We only mention here another fruitful consequence, the relation between linear series of points and rational mappings of an irreducible curve Γ into a projective space $\boldsymbol{P}_r(\boldsymbol{C})$: such a mapping can be written

Theme E

$$\phi : \zeta \to (P_1(\zeta), P_2(\zeta), \cdots, P_{r+1}(\zeta)),$$

where the P_j are homogeneous polynomials in the homogeneous coordinates of ζ,

all of the same degree: if Γ' is the image of Γ by ϕ, the points of intersection of Γ by the system of curves $\sum_j \lambda_j P_j = 0$ are the inverse images by ϕ of the points of intersection of Γ' by variable hyperplanes. This observation, in connection with the theory of linear series, enables one to choose the P_j in such a way that ϕ is a birational transformation and Γ' has *no singular points*. Furthermore, the curve Γ' having these properties is uniquely determined up to a birational and *bijective* transformation (one says it is *the* nonsingular *model* of the field of rational functions of Γ).

VI c: Integrals of differential forms on higher dimensional varieties. As soon as 1870, Cayley, Clebsch and M. Noether inaugurated the study of abelian integrals on irreducible algebraic surfaces, by considering, on a surface S in $\boldsymbol{P}_3(\boldsymbol{C})$ given by an equation $F(x, y, z) = 0$ in nonhomogeneous coordinates, double integrals of type $\iint R(x, y, z)dxdy$, where R is a rational function; after 1885, Picard began a thorough investigation of the properties of these integrals, as well as of simple integrals $\int P(x, y, z)dx + Q(x, y, z)\,dy$, where P, Q are rational and the differential is exact*. His method, which (conveniently generalized) is still very useful, consists in looking at the sections of the surface by the planes $y =$ const., applying Riemann's theory to abelian integrals on these curves (which in general are irreducible), and studying the way in which they depend on the parameter y; in particular, if p is the genus of the curve for general values of y, the $2p$ periods of the abelian integrals of the first kind satisfy a linear differential equation or order $2p$ (as functions of y), the so called Picard-Fuchs equation, which plays an important part in the theory. The algebraic surfaces considered by these mathematicians were usually supposed to be without singular points, or at most to have only "nice" singularities (double curves with distinct tangent planes except at finitely many points and no singular points except finitely many triple points); starting with M. Noether, many attempts were made to prove that any algebraic surface could be transformed into surfaces without singularities (not necessarily immersed in $\boldsymbol{P}_3(\boldsymbol{C})$,

Theme F

Theme C

* The exact meaning of a simple integral $\int P(x, y, z)\,dx$ consists in assigning to each piecewise differentiable mapping $t \to (x(t), y(t), z(t))$ of an interval $[a, b] \subset \boldsymbol{R}$ into S (a "singular 1-simplex") the number $\int_a^b P(x(t), y(t), z(t))\,x'(t)\,dt$. Similarly, the double integral $\iint R(x, y, z)\,dxdy$ assigns to each piecewise differentiable mapping $(u, v) \to (x(u, v), y(u, v), z(u,v))$ of a triangle $T \subset \boldsymbol{R}^2$ into S (a "singular 2-simplex") the number

$$\iint_T R(x(u, v),\, y(u, v),\, z(u, v))\, \frac{\partial(x, y)}{\partial(u, v)}\, dudv.$$

One can then define in an obvious way the value of simple (resp. double) integrals over 1-chains (resp. 2-chains), i.e., formal linear combinations of 1-simplices (resp. 2-simplices) with coefficients in $\boldsymbol{Z}$ (or in $\boldsymbol{R}$, or in $\boldsymbol{C}$). Generalizations to higher dimensions are obvious, once one defines an n-simplex as a piecewise differentiable mapping of the "standard n-simplex" defined by the inequalities $x_j \geqq 0$ $(1 \leqq j \leqq n)$, $x_1 + x_2 + \ldots + x_n \leqq 1$ in $\boldsymbol{R}^n$.

but in higher dimensional projective spaces), but no satisfactory proof was found until much later.

Very early it appeared that the theory of algebraic surfaces exhibited some features which had no counterpart in the theory of algebraic curves. Two irreducible surfaces without singularities may be birationally equivalent without being isomorphic. If p_g denotes the number of linearly independent double integrals of the first kind on an irreducible surface S (i.e., integrals which are finite over any 2-cell of S), the corresponding number for a surface S' birationally equivalent to S is not necessarily the same. The number p_g is the obvious counterpart of the genus of a curve; but very soon also, it was realized that the other definition of the genus of a curve, using the "adjoints" of Riemann, also generalized to surfaces, but might give a number p_a different from p_g (see in VIII-a its exact definition in modern terms); p_g was called the *geometric genus* and p_a the *arithmetic genus* of S, and the difference $q = p_g - p_a$ (which is always $\geqq 0$) the *irregularity* of the surface (for instance, Cayley found that for *ruled* surfaces $p_g = 0$ and $p_a < 0$ in general).

Theme A

It soon also became apparent that the properties of abelian integrals on a surface or a higher dimensional variety were to a large extent subordinate to the topological properties of the variety. H. Poincaré had particularly in mind the applications to algebraic geometry when, in 1895, he started to give mathematical substance to Riemann's intuition of higher dimensional "Betti numbers" by inventing the "simplicial" machinery which made rigorous proofs possible*; algebraic varieties (and more generally analytic varieties) are amenable to this technique due to the fact that they are *triangulable*, a fact for which Poincaré himself sketched a proof, which was later made entirely rigorous by van der Waerden. Using this machinery and the Picard technique of variable plane sections, Poincaré was able to bring to a satisfactory conclusion previous efforts by Picard and the Italian geometers and to prove that the irregularity q of an algebraic surface without singularity is equal to $R_1/2$, where R_1 is the first Betti number, and also equal to the number of independent simple abelian integrals of the first kind. Around 1920, Lefschetz considerably developed these techniques and generalized them to algebraic varieties of arbitrary dimension, concentrating in particular on the determination of the number of cycles on such a variety V which are homologous to cycles contained in algebraic subvarieties of V: for instance, if V is a projective variety of complex dimension n, and H a hyperplane section of V, the natural mappings

Theme F

* Let us recall that to an n-chain is attached a well determined $(n-1)$ - chain, its boundary; *n-cycles* are the n-chains whose boundary is 0, and the n-th *homology group* $H_n(M, \boldsymbol{Z})$ (resp. $H_n(M, \boldsymbol{R})$, resp. $H_n(M, \boldsymbol{C})$) of a manifold M, with coefficients in $\mathbf{Z}$ (resp. $\boldsymbol{R}$, $\boldsymbol{C}$) is the quotient of the group of n-cycles with coefficients in $\mathbf{Z}$ (resp. $\boldsymbol{R}, \boldsymbol{C}$) by the subgroup consisting of the boundaries of the $(n+1)$-chains. The Betti number R_p is the dimension of the real vector space $H_p(M, \boldsymbol{R})$.

$H_i(H,\boldsymbol{Z}) \to H_i(V,\boldsymbol{Z})$ of homology groups are bijective for $0 \leqq i \leqq n-2$ and surjective for $i = n-1$. He also showed that for an algebraic variety V, one had $R_{2p} > 0$, $R_p \geqq R_{p-2}$ for $p \leqq n$ (complex dimension of V) and that the Betti numbers R_{2p+1} of odd dimension were even.

VI d: Linear systems and the Italian school. The definition of divisors, given in VI-a, carries over to any field K finitely generated over $\boldsymbol{C}$; on a nonsingular model V having K as field of rational functions, the discrete nontrivial valuations of K now correspond to irreducible subvarieties of V of *codimension* 1. It is still true that $\deg((f)) = 0$ for principal divisors, and that $L(D)$ is a finite dimensional subspace of K for all divisors D. The concept of *linear system* of subvarieties of codimension 1 may therefore be associated to the notion of divisor as in VI-b. Around 1890, the Italian school of algebraic geometry, under the leadership of a trio of great geometers: Castelnuovo, Enriques and (slightly later) Severi, embarked upon a program of study of algebraic surfaces (and later higher dimensional varieties) generalizing the Brill-Noether approach *via* linear systems: they chiefly worked with purely geometric methods, such as projections or intersections of curves and surfaces in projective space, with as little use as possible of methods belonging either to analysis and topology, or to "abstract" algebra.

These limitations implied serious difficulties in the definition of the main concepts and the use of geometric methods. The chief trouble was that whereas on curves one can work almost exclusively with *positive* divisors, this is not the case any more for surfaces: for instance if $p_g = 0$, the canonical divisor (defined as in VI-a, but for meromorphic differential 2-forms) is not equivalent to a positive divisor, hence does not correspond to a linear system of curves. This compelled the Italians to introduce complicated "virtual" notions for linear systems, which obscured the significance of much of their results.

Working under such considerable handicaps, it is amazing to see how many new and deep results were discovered by the Italian geometers. It would be extremely long and intricate to describe these results in their own language (see for instance [**16**]) and we shall postpone the definition of the most important notions which they introduced until we can use the much simpler modern formulation.

Let us only mention here a few of the beautiful theorems characterizing (up to birational equivalence) simple types of surfaces by the values of the arithmetical genus p_a and new invariants defined by, Enriques, the *plurigenera* P_k $(k \geqq 2)$: a *rational* surface (i.e., birationally equivalent to a plane) is characterized by the relations $p_a = 0$, $P_2 = 0$, surfaces with $p_a < -1$ are ruled, whereas the surfaces such that $P_4 = P_6 = 0$ are either rational or ruled; finally, a surface for which $p_a = P_3 = 0$ and $P_2 = 1$ is birationally equivalent to the Enriques surface of degree 6 having the 6 edges of a tetrahedron as double lines (it is not a rational surface, although $p_g = 0$).

Theme A

VII. SIXTH PERIOD: "NEW STRUCTURES IN ALGEBRAIC GEOMETRY" (1920–1950)

The general trend towards the unification of mathematics by the study of the *structures* underlying each theory, which started to get momentum in the 1920's, was particularly apparent in the development of algebraic geometry; the striking kinships between algebraic varieties and complex manifolds on the one hand, algebraic numbers on the other, which had been discovered in earlier periods, now became organic parts of the fundamental concepts of algebraic geometry. One of the effects of this broadened point of view was to loosen the exclusive grip held until then by projective and birational methods over algebraic geometry, and prepare the way for a far more flexible approach.

VII a: Kählerian varieties and the return to Riemann. Ever since Gauss's fundamental paper of 1826 on the theory of surfaces and Riemann's inaugural lecture of 1854 defining n-dimensional riemannian geometry, the concept of *differential manifold*, defined by "maps" and differentiable "transition functions" between maps*, had gradually become more and more precise as the fundamental topological concepts needed to express them were defined and studied in the last part of the 19th century and the beginning of the 20th. One of the most important developments in that direction was the introduction of the general concept of *exterior differential p-form* on a differential manifold (locally defined by expressions

$$\sum_{i_1 < i_2 < \cdots < i_p} A_{i_1 i_2 \cdots i_p}(x) dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}$$

in the local coordinates) and of their integrals on *p-chains* (generalizing the earlier notions of "curvilinear" and "surface" integrals), due to H. Poincaré and E. Cartan. At the very beginning of his papers on algebraic topology, Poincaré had pointed out the connection between the homology of a compact differential manifold V and the exterior differential forms on V (of which the classical Stokes' theorem is the simplest example). This was made precise by De Rham's famous theorems in 1931, starting from the duality between chains and forms given by the integral $\langle C, \omega \rangle = \int_C \omega$; due to the generalized Stokes' formula $\langle C, d\omega \rangle = \langle bC, \omega \rangle$ (where b is the boundary and d the exterior derivative), this yields a duality, pairing the real homology groups $H_i(V, \boldsymbol{R})$ of V and the cohomology groups $H^i(\Lambda)$**, where Λ is the "complex" of exterior differential forms

$$0 \xrightarrow{d} \Lambda^1 \xrightarrow{d} \Lambda^2 \xrightarrow{d} \cdots \xrightarrow{d} \Lambda^n \xrightarrow{d} 0 \qquad (n = \dim V), \tag{4}$$

(Λ^j is the $\boldsymbol{R}$-vector space of the j-forms).

* If M is a differential manifold of dimension $n, \phi: U \to \mathbf{R}^n$, $\psi: Y \to \mathbf{R}^n$ two maps of open sets U, V, of M onto $\mathbf{R}^n$, the "transition function" from U to V is the mapping (only defined when $U \cap V \neq \emptyset$) $x \mapsto \psi(\phi^{-1}(x))$ of $\phi(U \cap V)$ onto $\psi(U \cap V)$.

** $H^i(\Lambda)$ is the quotient of the kernel $d^{-1}(\Lambda^{i+1})$ by the image $d(\Lambda^{i-1})$.

A projective algebraic variety without singularity of (complex) dimension n has a natural underlying structure of differential manifold of dimension $2n$, but in fact it has a much richer structure. In the first place, it is a *complex* manifold, which means that for the "maps" which define the differential structure and which take their values in $\boldsymbol{C}^n$ ($= \boldsymbol{R}^{2n}$), the "transition functtions" are *holomorphic*; it follows that the space $\Lambda_{\boldsymbol{C}}^p$ of (complex) differential p-forms for $1 \leqq p \leqq 2n$ decomposes naturally into a direct sum of vector spaces $\Lambda_{\boldsymbol{C}}^{r,s}$ corresponding to the pairs of integers such that $r + s = p$; for $r \leqq n$ and $s \leqq n$, the forms in $\Lambda_{\boldsymbol{C}}^{r,s}$ (called forms of *type* (r, s)) are those which for *complex* local coordinates $z^1, z^2, \cdots, z^n$, are written

Theme F

$$\sum A_{j_1 \cdots j_r k_1 \cdots k_s}(x) dz^{j_1} \wedge \cdots \wedge dz^{j_r} \wedge d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_s}, \tag{5}$$

where the $A_{j_1 \cdots k_s}$ are differentiable functions with complex values (not holomorphic in general). For $r > n$ or $s > n$, one takes $\Lambda_{\boldsymbol{C}}^{r,s}$ as reduced to 0 by convention.

But this is not the end of the story. It is possible to define on a projective complex space, and by restriction on any complex compact submanifold of such a space (which is necessarily an algebraic variety by a theorem of Chow) a riemannian ds^2 which is *kählerian*, i.e., can be written locally as a hermitian form

$$ds^2 = \sum_{j,k} h_{jk} dz^j d\bar{z}^k \text{ with } h_{kj} = \bar{h}_{jk}$$

which has the property that the corresponding exterior 2-form (which is *real* valued)

$$\Omega = (i/2) \sum_{h,j} h_{jk} d\bar{z}^k \wedge dz^j \tag{6}$$

is *exact* (i.e., $d\Omega = 0$).

Beginning around 1930, Hodge, in a series of remarkably original papers, showed how to use these facts to investigate the homology of compact kählerian varieties. On a riemannian manifold, Beltrami had shown that it is possible to define an operator which generalizes the usual laplacian, and therefore enables one to define *harmonic functions* on the manifold. By a very imaginative generalization, Hodge was able to define similarly, on any compact riemannian manifold, the notion of *harmonic exterior differential forms*, and to prove that there existed a unique such form in any cohomology class in any $H^j(\Lambda)$; from that result, he deduced the uniqueness and existence of a harmonic p-form having given periods on homologically independent p-cycles, thus obtaining a complete generalization of Riemann's fundamental result, and showing that Riemann's use of "Dirichlet's principle" was far more than a technical device (fortunately for Hodge, the theory of elliptic partial differential equations had advanced far enough to spare him the difficulties which had plagued Riemann's approach). Turning next to complex kählerian manifolds, the space H^p of harmonic p-forms with complex coefficients splits into a direct sum of $p + 1$ spaces $H^{r,s}$ consisting of (complex) harmonic forms

of type (5), for $r+s=p$ (with $H^{r,s}=0$ if $r>n$ or $s>n$); it can be shown that $H^{p,0}$ consists exactly of the *holomorphic* p-forms (or "differential forms of the first kind"), i.e., those for which in (5), $s=0$ and the $A_{j_1\cdots j_p}$ are holomorphic. As complex conjugation transforms $H^{r,s}$ into $H^{s,r}$, they have the same dimension, and this shows that the dimension of H^p, i.e., the Betti number R_p, is even when p is odd. On the other hand, one easily verifies that the (real) 2-form Ω defined in (6) is harmonic, as well as all its exterior powers, which proves that $R_{2k}\geqq 1$ for every integer k. Finally, $\phi\mapsto\Omega\wedge\phi$ is shown to be an injective mapping of H^p into H^{p+2} for $p\geqq n-2$, from which the inequality $R_{p+2}-R_p\geqq 0$ follows; all the Lefschetz's theorems on Betti numbers of algebraic varieties are thus "explained" and shown to belong in fact to the theory of kählerian manifolds (there are compact kählerian manifolds which are not isomorphic to projective algebraic varieties). We shall return to the Hodge's theory when in the next period it merges into sheaf cohomology.

VII b: Abstract algebraic geometry. It is well known that, from 1900 to 1930, the general concepts of algebra (mostly confined until then to real or complex numbers) were developed in a completely abstract setting, the notion of algebraic *structure* (such as group, ring, field, module, etc.) becoming the fundamental one and relegating to second place the nature of the mathematical *objects* on which the structure was defined. It was therefore quite natural to think of an "abstract" extension of algebraic geometry, in which the coefficients of the equations and the coordinates of the points would belong to an arbitrary field. Already Dedekind and Weber, in their 1882 paper, had observed that all their arguments only used the fact that the basic field was algebraically closed (and of characteristic 0, a notion which had not yet been defined then). Even notions which seem linked to analysis, such as derivatives and differentials, had algebraic counterparts: a *derivation* in a commutative ring A is an additive mapping $x\mapsto Dx$ of A into itself such that $D(xy)=x\cdot Dy+(Dx)\cdot y$, and a *differential* is an A-linear mapping $\omega\colon\mathfrak{D}\to A$ of the A-module of all derivations into A; for each $x\in A$, dx is the linear form $D\mapsto Dx$ on $\mathfrak{D}$, and p-forms are defined by the usual methods of exterior algebra.

Theme G

The motivation for the development of abstract algebraic geometry was therefore a natural outcome of the progress of algebra; after 1930, a more powerful impulse was to come from number theory, as we shall see below.

As it was apparent that a large part of the foundations of classical algebraic geometry came from geometric intuition, more or less justified by appeals to analysis or topology, a thorough examination of the basic concepts, from the exclusive viewpoint of algebra, was necessary in order to carry out an ambitious program of algebraic geometry over an arbitrary field. This groundwork, which at the same time created most of modern commutative algebra, was chiefly due to E. Noether, W. Krull, van der Waerden, and F. K. Schmidt in the period 1920–1940, and to Zariski and A. Weil from 1940 on.

The first two of these mathematicians use the geometric language very sparsely; their results are almost always expressed in the language of rings and ideals, and it was only after 1940 that the importance of their work was properly appreciated: the decomposition into primary ideals in noetherian rings, the properties of integrally closed rings, the extensive use of valuations, the notion of localization and the fundamental properties of local rings are all due to them. (A local ring is a commutative ring A in which there is only one maximal ideal. The typical example consists of the rational functions (elements of the field $\boldsymbol{C}(X)$) for which a given point $\zeta \in \boldsymbol{C}$ is not a pole: they form the local ring of $\boldsymbol{C}(X)$ at the point ζ). A similar remark may be made on the foundational work of Zariski, probably the deepest one in that period; although it is usually expressed in the language of projective geometry, it mostly belongs to local algebra and its central position in algebraic geometry was only recognized in the next period. The contribution of F. K. Schmidt (in connection with his work on number theory which we describe below) essentially consisted in extending the Dedekind-Weber theory to curves defined over an algebraically closed field of *any characteristic*.

The most conspicuous progress realized during that period is the successful definition, in algebraic geometry over an arbitrary field, of the concepts of *generic point* and of *intersection multiplicity*, due to the combined efforts of van der Waerden and A. Weil. The Italians (not to speak of their predecessors) used these notions with a freedom which, to their critics of the orthodox algebraic school, bordered on recklessness. As long as the underlying field was $\boldsymbol{C}$, the notion of "elements in general position" could be easily justified by an appeal to continuity (although the Italians seldom bothered to prove that these elements formed *open sets* in the spaces they considered). On the other hand, Lefschetz had made the elementary but fundamental observation that when two subvarieties U, V of $\boldsymbol{P}_n(\boldsymbol{C})$, of complementary dimensions r and $n - r$, intersect transversally in simple points, the number of these points is equal (for convenient orientations) to the *intersection number* $(U \cdot V)$ of the *cycles* U, V, in the sense of algebraic topology; as this number is known to be invariant under homology, it was quite natural to take it as the number of intersections of U and V (counted with multiplicities) in the most general cases. This justified the extensive use of intersection multiplicity by the Italian geometers, in particular the "self-intersection" number $(C \cdot C)$ of a curve on an algebraic surface. (Unfortunately, the complexity of the Italian definitions was such that it was often impossible to be sure that the same words meant the same things in two different papers; hence the numerous controversies between geometers of that school, such as the one which occurred as late as 1943 between Enriques and Severi, see **[4]** and **[10]**.)

These foundations of course disappeared in algebraic geometry over an arbitrary field, and this was one of the reasons why no algebraic proofs valid over any field (even of characteristic 0) had been found for the results obtained in the theory of algebraic surfaces by transcendental or geometric methods. In 1926, van der Waerden

saw that to gain the freedom which Analysis gave for classical geometry over the complex field, one had only to return to the process which had allowed the passage from real to complex geometry, namely *enlarge* the field k to which the coefficients of the equations of a variety and the coordinates of its points are supposed to belong: if K is any extension of k, these equations are still meaningful when the coordinates are taken in K. Giving a general form to ideas which went back at least to Gauss, he introduced the idea of *specialization* over k of any set of elements $x_1, \cdots, x_m$ in an arbitrary extension K of k: it is a mapping which to each x_j assigns an element x'_j of an extension K' of k (which may be equal to K), in such a way that for every homogeneous polynomial $P \in k[X_1, \cdots, X_m]$ for which $P(x_1, \cdots, x_m) = 0$, one also has $P(x'_1, \cdots, x'_m) = 0$ (van der Waerden always works in projective spaces, or finite products of such spaces). Suppose then that V is an irreducible algebraic variety in $\boldsymbol{P}_n(k)$, and let K be the field of rational functions on V; one may assume that V is not contained in a hyperplane of $\boldsymbol{P}_n(k)$; for $1 \leqq j \leqq n$, the restriction ξ_j to V of the rational function $x \mapsto x^j/x^0$ (where $x^0, x^1, \cdots, x^n$ are homogeneous coordinates of a point, $x \in \boldsymbol{P}_n(k)$) is an element of K; if V_K is the variety in $\boldsymbol{P}_n(K)$ defined by the same equations as V, the point $(1, \xi_1, \cdots, \xi_n)$ belongs to V_K. Van der Waerden calls this point a *generic point* of V, for it is immediate to check that for *any* extension K' of k, *any* point of $V_{K'}$ is a specialization of $(1, \xi_1, \cdots, \xi_n)$. Such points can then be used in the same way as the "general points" of the Italians, despite their apparently tautological character: any theorem proved for generic points (and of course expressible by algebraic *equations* (not inequalities!) between their coordinates) is valid for *arbitrary* points of corresponding varieties. Van der Waerden then proceeded to apply this new tool with great virtuosity to many problems of algebraic geometry, and in particular to the definition of multiplicity of intersection of two varieties in abstract algebraic geometry, which had not yet been given a meaning except in the case of the intersection of two curves on a surface without singularity. However, Poncelet, as a consequence of his general vague "principle of continuity," had already proposed to define the intersection multiplicity at one point of two subvarieties U, V of complementary dimensions by having V (for instance) *vary* continuously in such a way that for some position V' all the intersection points with U should be *simple*, and counting the number of these points which collapsed to the given point when V' tended to V; in such a way, the total number of intersections (counted with multiplicities) would remain constant ("principle of the conservation of number"), and it is thus that Poncelet proved Bézout's theorem, by observing that a curve C in the plane belonged to the continuous family of all curves of the same degree m, and that in that family there existed curves which degenerated into a system of straight lines, each meeting a fixed curve Γ of degree n in n distinct points. Many mathematicians in the 19th century had extensively used such arguments, and in 1912, Severi had convincingly

Theme D

Theme C

argued for their essential correctness. The concept introduced by van der Waerden was based on similar ideas: under suitable conditions, the multiplicity of a solution $y = (y_0, \cdots, y_n) \in \boldsymbol{P}_n(k)$ of a system of equations $P_\mu(x, y) = 0$, where x is a point of an irreducible variety V, is the number of the solutions η of the system $P_\mu(\xi, \eta) = 0$, where ξ is the generic point of V, which specialize to y when ξ specializes to x. Using this definition, he was finally able to attach to every irreducible component C of the intersection of two irreducible varieties V, W of an "ambient" nonsingular variety U, an integer $i(C, V \cdot W;\ U) \geqq 0$, the multiplicity of C in $V \cap W$, provided *all* irreducible components of $V \cap W$ were "proper," i.e., had a dimension equal to $\dim V + \dim W - \dim U$.

Unfortunately, this restriction considerably reduced the usefulness of the notion of multiplicity. Using more powerful algebraic devices, A. Weil could define an intersection multiplicity $i(C, V \cdot W;\ U)$ when it is *only* supposed that C is proper (the other components of $V \cap W$ can have larger dimensions); furthermore, he showed that this number did not depend on the method used to define it (other, quite different methods, were later given by Chevalley and Samuel), once it possessed the "natural" properties similar to those of the intersection number in algebraic topology; this he showed to be the case for his definition, and it enabled him to develop in abstract algebraic geometry a calculus of "cycles" patterned on the calculus of chains introduced by Poincaré (irreducible subvarieties replacing simplices). In this context, divisors on an irreducible variety of dimension n were the cycles of dimension $n - 1$ (one also says that they have *codimension* 1).

Weil then went on to break away, for the first time, from projective algebraic geometry: for his purposes (see below) he needed constructions of algebraic varieties similar to the "gluing together" constructions of manifolds in algebraic topology or differential geometry, which had been familiar since the beginning of the century; he showed that this could be done by using as "transition functions" biregular mappings of complements of subvarieties in affine varieties (the Zariski topology was not yet in use at that time), and he could also define in this context the notion of "complete variety" which is the counterpart of the concept of compact space in "abstract" algebraic geometry (in classical projective geometry, all algebraic subvarieties are complete).

Theme E

VII c: Zeta functions and correspondences. A. Weil's work was chiefly motivated by problems which had arisen in the early 1920's in number theory. In his thesis of 1923, E. Artin had observed that algebraic congruences modulo a prime p, in 2 variables, i.e., of the form $F(x, y) \equiv 0 \pmod{p}$, where F is a polynomial with integral coefficients, could be interpreted as algebraic equations over the prime field $\boldsymbol{F}_p = \boldsymbol{Z}/p\boldsymbol{Z}$ (and similarly the "higher congruences" in the sense of Dedekind were algebraic equations over an arbitrary *finite* field $\boldsymbol{F}_q$ $(q = p^d)$). He further noticed that the analogy, already exploited by Dedekind and Weber, of finite extensions of

the field $C(X)$ with algebraic number fields, was here much closer, since the residual fields of the valuations of a finite extension K of $F_q(X)$ are *finite* fields (extensions of F_q) just as for number fields (whereas they are equal to C in classical algebraic geometry). This enabled him to define, in complete analogy with the Riemann-Dedekind zeta function of an algebraic number field, the *zeta function of* K, and to extend to it the classical theory: functional equation and the location of the poles. However, his treatment was entirely algebraic, without any kind of geometric interpretation; a little later, F. K. Schmidt observed that a much simpler and more natural treatment was achieved if one completely modeled the theory after Dedekind and Weber, by introducing *divisors* (or "points of the abstract Riemann surface") instead of ideals; it can then easily be shown that the zeta function can be defined by the equation (for $u = q^s$)

$$\frac{d}{du}(\log Z(u)) = \sum_{m=1}^{\infty} N_m u^{m-1}, \qquad Z(0) = 1,$$

where N_m is the number of points of the curve whose coordinates belong to the extension F_{q^m} of F_q of degree m. It turns out that this function is much simpler than in the classical case; in fact it is a rational function

$$Z(u) = P_{2g}(u)/(1-u)(1-qu),$$

where P_{2g} is a polynomial of degree $2g$ (g being the genus of K). F. K. Schmidt further discovered the remarkable fact that the functional equation

$$Z(1/qu) = q^{1-g}u^{2-2g}Z(u)$$

was nothing else but the analytic expression of the Riemann-Roch theorem!

At the same time, arithmeticians had been endeavoring to obtain an evaluation of N_1, the number of points of the nonsingular curve Γ corresponding to K with coordinates in F_q, and had obtained estimates of the form $|N_1 - (q+1)| \leqq Cq^\alpha$, with C independent of q and $1/2 < \alpha < 1$; they had observed that $\alpha = 1/2$ would be the best possible result. Hasse became interested in the problem and remarked that the result was a consequence of the so-called "Riemann hypothesis for curves over finite fields," namely the fact that all the zeroes of the polynomial P_{2g} lay on the circle $|u| = q^{1/2}$, this fact implying the inequality

$$|N_1 - (q+1)| \leqq 2g \cdot q^{1/2} \tag{7}$$

in an elementary way. In 1934, he succeeded in proving this result for $g = 1$, by adapting to the case of finite fields ideas from the theory of complex multiplication of elliptic functions. He and Deuring observed furthermore that an extension to values $g \geqq 2$ would have to be based on the theory of correspondences.

This is what A. Weil proceeded to do. An irreducible correspondence between two irreducible curves Γ_1, Γ_2 is an irreducible curve on the surface $\Gamma_1 \times \Gamma_2$, and in general a *correspondence* between Γ_1 and Γ_2 is a *divisor* on $\Gamma_1 \times \Gamma_2$; degenerate correspondences are those of the form $\{x_1\} \times \Gamma_2$ or $\Gamma_1 \times \{x_2\}$ $(x_i \in \Gamma_i)$ and linear combinations of such with integral coefficients; correspondences are called *equivalent* if they differ by the sum of a principal divisor and a degenerate correspondence. For $\Gamma_1 = \Gamma_2 = \Gamma$, one defines as in set theory the *composition* $X \circ Y$ of two correspondences; it can be proved that, together with the addition of divisors, this defines on the set of equivalence classes $\mathfrak{A}(\Gamma)$ a structure of *ring* with unit element (the class of the diagonal Δ of $\Gamma \times \Gamma$). The degrees $d(X)$ and $d'(X)$ of a correspondence are defined as the integers, such that the first (resp. second) projection of X is the cycle $d(X) \cdot \Gamma$ (resp. $d'(X) \cdot \Gamma$); on the other hand, for two correspondences X, Y which intersect properly, $I(X \cdot Y)$ is the degree of the cycle $X \cdot Y$. One can then show that the integer

Theme B

$$S(X) = d(X) + d'(X) - I(X \cdot \Delta)$$

only depends on the equivalence class ξ of X, and has the property of a *trace*, i.e., $S(\xi \cdot \eta) = S(\eta \cdot \xi)$ for two elements of $\mathfrak{A}$. Furthermore, to each correspondence X is associated another one X', deduced from X by the symmetry automorphism of $\Gamma \times \Gamma$; if ξ, ξ' are the classes of X and X', one has $S(\xi \cdot \xi') \geqq 0$, equality being only possible for $\xi = 0$ in $\mathfrak{A}$. This theory was first developed in 1885 by Hurwitz, using Riemann's theory of abelian integrals, and the inequality for the trace was obtained by Castelnuovo (of course for the classical case); using his theory of intersection multiplicities, A. Weil was able to extend all these results to curves over arbitrary fields. He then observed that in the Hasse problem, the number N_m was exactly $I(F^m \cdot \Delta)$, where F is the "Frobenius correspondence" which to each point of Γ associates its transform by the automorphism of Γ corresponding to the automorphism $t \mapsto t^q$ of the algebraic closure of $\boldsymbol{F}_q$; from which it follows by definition that $S(F^m) = 1 + q^m - N_m$, and expressing the inequality $S(\xi \cdot \xi') \geqq 0$ where ξ is the class of $a \cdot \Delta + b \cdot F^m$, for arbitrary integers a, b, one gets $|N_m - q^m - 1| \leqq 2g \cdot q^{m/2}$, which generalizes (7) and implies the "Riemann hypothesis."

VII d: Equivalence of divisors and abelian varieties. The introduction of varieties of arbitrary dimension had been particularly useful because it allowed to consider as points in a projective space of sufficiently high dimension geometric objects such as lines, conics, etc. In 1937, Chow and van der Waerden showed quite generally that it is possible to consider the irreducible algebraic subvarieties of given dimension and degree in a given $\boldsymbol{P}_n(k)$ as the points of some algebraic variety in a suitable $\boldsymbol{P}_N(k)$. From this result it follows that it is possible to give a precise meaning (for an arbitrary field k) to the concepts of "specialization of cycles" and of "algebraic family of cycles" which had been used in the classical case by the Italian school.

In particular, one can define the concept of *algebraic equivalence* of two divisors D_1, D_2 on a nonsingular variety V as meaning that they belong to a common irreducible algebraic family of divisors. Another concept of equivalence is *numerical equivalence*, meaning that for any curve C on V, the intersection numbers $(D_1 \cdot C)$ and $(D_2 \cdot C)$ are equal. If one denotes by G, G_n, G_a, G_l the group of divisors on V and its subgroups formed of divisors equivalent to 0 for numerical, algebraic and linear equivalence, one has $G \supset G_n \supset G_a \supset G_l$. Severi for the classical case, and Matsusaka for arbitrary characteristic proved that the group G_n/G_a is finite. A deeper result, proved by Severi for complex algebraic surfaces, following earlier results of Picard, is that the group G/G_n is a free finitely generated commutative group $\boldsymbol{Z}^p$; this result was extended by Néron for arbitrary fields and in any dimension. Finally, it was known since Riemann that for an irreducible algebraic curve over $\boldsymbol{C}$, the group G_a/G_l was naturally endowed with a structure of g-dimensional algebraic nonsingular variety (g being the genus of the curve) which, as a topological group, is isomorphic to a *complex torus* $\boldsymbol{C}^g/\Gamma$, where Γ is a lattice in $\boldsymbol{C}^g$ (discrete group isomorphic to $\boldsymbol{Z}^{2g}$); this variety is called the *Jacobian* of the curve, and it had been used since Clebsch to study the geometry on an algebraic curve. In general, a complex torus $\boldsymbol{C}^n/\Gamma$, where Γ is a lattice in $\boldsymbol{C}^n$ (isomorphic to $\boldsymbol{Z}^{2n}$) can only be given the structure of an algebraic variety if the lattice Γ satisfies certain bilinear relations which had been already found by Riemann; it is then called an *abelian variety*. The work of Picard and his successors proved that for an arbitrary nonsingular algebraic variety V over $\boldsymbol{C}$, the group G_a/G_l was again equipped with a structure of abelian variety, called the *Picard variety* of V. Following his work on the Riemann hypothesis, A. Weil developed the general theory of abelian varieties over an arbitrary field (as "abstract" varieties), and was able to define the Jacobian of a curve. Later work of Chow and Matsusaka proved that abelian varieties can still be imbedded in projective space in the general case, and extended to any field the definition of the Picard variety.

VIII. SEVENTH PERIOD: "SHEAVES AND SCHEMES" (1950–)

After 1945, the considerable progress brought in algebraic topology, differential topology and the theory of complex manifolds by the introduction of sheaves and spectral sequences (both due to J. Leray) completely renewed the concepts and methods of algebraic geometry, both "classical" and "abstract," simplifying old definitions and results and opening new ways leading to the solution of old problems.

VIII a: The Riemann-Roch theorem for higher dimensional varieties and sheaf cohomology. The Riemann-Roch problem for an irreducible algebraic variety V is the computation of the dimension $l(D)$ of the vector space $L(D)$ for an arbitrary divisor D on V by some formula similar to the Riemann-Roch theorem for curves (3).

The Italian geometers had attacked the problem for surfaces, but succeeded only in getting a *lower bound* for $l(D)$, expressed in terms of $\deg(D)$ and birational invariants of the surface S, of D and of $\Delta - D$ (where Δ is a canonical divisor).

Theme F

In the 1930's, study of differential geometry and in particular of E. Cartan's method of moving frames had finally led to the definition of *vector bundles* over a differential manifold M: such a bundle is a differential manifold $\mathbf{E}$ with a projection $p\colon \mathbf{E} \to M$ such that the fibers $p^{-1}(x)$ for any $x \in M$ are real (resp. complex) vector spaces of fixed dimension r (the *rank* of $\mathbf{E}$), and locally on M, $\mathbf{E}$ looks like the product of M and $\boldsymbol{R}^r$ (resp. $\boldsymbol{C}^r$); in other words each point of M has an open neighborhood U for which there is a diffeomorphism ϕ transforming $p^{-1}(U)$ onto $U \times \boldsymbol{R}^r$ (resp. $U \times \boldsymbol{C}^r$) in such a way that ϕ transforms *linearly* each fiber $p^{-1}(x)$ into $\{x\} \times \boldsymbol{R}^r$ (resp. $\{x\} \times \boldsymbol{C}^r$). A *section* of $\mathbf{E}$ is a differentiable mapping $s\colon x \mapsto s(x)$ of M into $\mathbf{E}$ such that $s(x) \in p^{-1}(x)$ for every $x \in M$. Over a complex manifold M, one can similarly define holomorphic vector bundles by taking $\mathbf{E}$ as a complex manifold, the projection p being holomorphic, the fibers $p^{-1}(x)$ complex vector spaces, and ϕ (in the above definition) being also holomorphic. Important examples of vector bundles are the *tangent bundle* $\mathbf{T}(M)$, where the fiber $p^{-1}(x)$ consists of the tangent vectors to M at x (so that the rank is $\dim(M)$), and the *bundle of p-covectors* on M, whose sections are the exterior differential p-forms on M (see VII a).

The concept of *divisor* can be generalized to arbitrary complex manifolds M: if (U_α) is an open covering of M, one considers in each U_α a meromorphic function h_α, such that in $U_\alpha \cap U_\beta$, h_β/h_α is holomorphic and $\neq 0$ everywhere; two such systems (h_α), (h'_λ) corresponding to coverings (U_α), (U'_λ) are identified if h_α/h'_λ is holomorphic and $\neq 0$ in $U_\alpha \cap U'_\lambda$ for any pair (α, λ) of indices, and these classes of systems (h_α) are called divisors on M. One sees that for projective algebraic varieties over $\boldsymbol{C}$, this notion coincides with the old one: for instance, if $M = \boldsymbol{P}_n(\boldsymbol{C})$, and $D = \sum_k m_k S_k$ is a divisor on M, where each S_k is an irreducible hypersurface defined by an equation $F_k(x_0, x_1, \cdots, x_n) = 0$, F_k being an irreducible homogeneous polynomial of degree d_k, one covers $\boldsymbol{P}_n(\boldsymbol{C})$ with the $n+1$ open sets U_j $(0 \leqq j \leqq n)$, U_j being defined by the relation $x_j \neq 0$; one can then take as meromorphic function h_j in U_j the function

$$x \mapsto x_j^{-d} \prod_k (F_k(x_0, \cdots, x_n))^{m_k}$$

with $d = \sum_k m_k d_k$. In 1950, A. Weil observed that to a divisor D on a complex manifold M was naturally attached a complex vector bundle of rank 1 (what one calls a *line bundle*) $\mathbf{B}(D)$: with the previous notations, one "glues together" the complex manifolds $U_\alpha \times \boldsymbol{C}$ by taking as "transition function" from U_α to U_β the function $(x, z) \mapsto (x, (h_\beta(x)/h_\alpha(x))z)$, holomorphic in $(U_\alpha \cap U_\beta) \times \boldsymbol{C}$. Furthermore, if s is a holomorphic section of $\mathbf{B}(D)$, the restrictions s_α of s to U_α are such that in $U_\alpha \cap U_\beta$ one has $s_\beta = (h_\beta/h_\alpha)s_\alpha$, hence there is a meromorphic function f on M such that

the restriction of f to U_α is s_α/h_α for each α; for an algebraic variety M this is equivalent to $(f) + D \geqq 0$, and therefore $\mathrm{L}(D)$ can be interpreted as the vector space $\Gamma(\mathbf{B}(D))$ of all *holomorphic sections* of the line bundle $\mathbf{B}(D)$. For instance, if $M = \boldsymbol{P}_n(\boldsymbol{C})$, and $D = H$, a hyperplane in $\boldsymbol{P}_n(\boldsymbol{C})$, the transition functions for $\mathbf{B}(H)$ are

$$(x, z) \mapsto \left(x, \frac{x_k}{x_j} z\right)$$

in $U_j \cap U_k$ (with the notations introduced above), and $\Gamma(\mathbf{B}(H))$ is the vector space of all linear forms $(x_0, \cdots, x_n) \mapsto \lambda_0 x_0 + \cdots + \lambda_n x_n$ in $\boldsymbol{C}^{n+1}$.

Now to each complex vector bundle $\mathbf{E}$ over a differential manifold M of dimension n are attached, for each even integer $2j \leqq n$, well determined elements $c_j(\mathbf{E})$ of the cohomology group $H^{2j}(M, \boldsymbol{Z})$ called the *Chern classes* of $\mathbf{E}$*; when M is a complex manifold of real dimension $2n$, the Chern classes of $\mathbf{T}(M)$ are simply written c_j $(1 \leqq j \leqq n)$ and called the *Chern classes of* M; the number $\langle c_n, M\rangle$ (where M is considered as $2n$-cycle) is the Euler-Poincaré characteristic

$$\chi(M) = \sum_{j=0}^{2n} (-1)^j R_j .$$

Using the interpretation of divisors by line bundles and Hodge's theory of harmonic forms, Kodaira was able in 1951 to obtain, for compact kählerian manifolds of complex dimension 2, a "Riemann-Roch formula" in which the missing terms from the formula found by the Italian geometers were expressed by means of Chern classes; in 1952 he found a similar formula for kählerian manifolds of dimension 3.

Meanwhile, H. Cartan and Serre had discovered that Leray's concept of *sheaf* led to a remarkably simple and suggestive expression of the main results of the theory of complex manifolds. The holomorphic functions in open sets of such a manifold M satisfy Leray's axioms: if $\mathcal{O}(U)$ is the set of the complex functions holomorphic in the open set $U \subset M$, then, for every open covering (V_α) of U, a function $f \in \mathcal{O}(U)$ is entirely determined by its restrictions $f \mid V_\alpha \in \mathcal{O}(V_\alpha)$, and conversely, given for each α an $f_\alpha \in \mathcal{O}(V_\alpha)$ such that f_α and f_β have the same restriction to $V_\alpha \cap V_\beta$ for all pairs (α, β), there exists an $f \in \mathcal{O}(U)$ such that $f \mid V_\alpha = f_\alpha$ for all α.

---

* One can define the concept of *direct sum* of vector bundles over M by defining it locally in an obvious way; for any differentiable map $f: M' \to M$, one defines the "pullback" $f^*(\mathbf{E})$ of a vector bundle $\mathbf{E}$ over M as the submanifold of the product $M' \times \mathbf{E}$ consisting of the pairs (x', z) such that $f(x') = p(z)$. The Chern classes of $\mathbf{E}$ can then be characterized by the following conditions, where one writes $c(\mathbf{E})$ for the sum $\sum_{j=0}^{\infty} c_j(\mathbf{E})$ (the sum is finite since the groups $H^{2j}(M)$ are 0 for $2j > \dim M$; one writes by convention $c_0(\mathbf{E}) = 1$): (i) $c(f^*(\mathbf{E})) = f^*(c(\mathbf{E}))$, where on the right hand side f^*: $H^*(M, \boldsymbol{Z}) \to H^*(M', \boldsymbol{Z})$ is the natural mapping deduced from $f: M' \to M$.

(ii) $c(\mathbf{E}_1 \oplus \mathbf{E}_2 \oplus \cdots \oplus \mathbf{E}_m) = c(\mathbf{E}_1) c(\mathbf{E}_2) \ldots c(\mathbf{E}_m)$ for any direct sum of vector bundles $\mathbf{E}_j$ over M^1 (product taken in the cohomology ring $H^*(M, \boldsymbol{Z})$).

(iii) $c(\mathbf{B}(H)) = 1 + h_n$ for a hyperplane $H \subset \boldsymbol{P}_n(\boldsymbol{C})$, $h_n \in H^2(\boldsymbol{P}_n(\boldsymbol{C}), \boldsymbol{Z})$ being the cohomology class orresponding to the homology class of the $(2n-2)$-cycle H by Poincaré duality.

The sheaf thus defined is called the structural sheaf of M and written $\mathcal{O}_M$; one writes $H^0(U, \mathcal{O}_M)$ instead of $\mathcal{O}(U)$. More generally, for any complex vector bundle $\mathbf{E}$ over M, one defines the sheaf $\mathcal{O}(\mathbf{E})$ by replacing $\mathcal{O}(U)$ by the set of sections $\Gamma(U, \mathbf{E})$ of $\mathbf{E}$ above U, written $H^0(U, \mathcal{O}(\mathbf{E}))$; in particular one writes Ω_X^p the sheaf corresponding to the complex bundle of p-covectors on M, so that $H^0(U, \Omega_X^p)$ is the set of holomorphic exterior differential p-forms on U; for a divisor D on M, one writes $\mathcal{O}_X(D)$ instead of $\mathcal{O}(\mathbf{B}(D))$.

There are many types of sheaves other than those derived from vector bundles, and the usefulness of sheaves derives from this versatility and from the many operations one can do with sheaves. In the first place, to a sheaf of groups $\mathcal{F}$ over M and to each point $x \in M$ is associated a group, the *stalk* $\mathcal{F}_x$ of $\mathcal{F}$ at x: for $\mathcal{O}(\mathbf{E}), \mathcal{O}(\mathbf{E})_x$ consists of the equivalence classes of sections of $\mathbf{E}$ over neighborhoods of x for the following relation: two sections are equivalent if they coincide on a neighborhood of x ("germs of sections"); the general definition of $\mathcal{F}_x$ is similar. For a sheaf of abelian groups $\mathcal{G}$ and a sheaf $\mathcal{N} \subset \mathcal{G}$ such that $\mathcal{N}_x$ is a subgroup of $\mathcal{G}_x$ for each x, one can then define a quotient sheaf $\mathcal{G}/\mathcal{N}$ such that $(\mathcal{G}/\mathcal{N})_x = \mathcal{G}_x/\mathcal{N}_x$. Each stalk $(\mathcal{O}_X)_x$ (written $\mathcal{O}_x$) is a local ring, and if $\mathcal{F}, \mathcal{G}$ are two sheaves such that $\mathcal{F}_x$ and $\mathcal{G}_x$ are $\mathcal{O}_x$-modules, then one can define a sheaf $\mathcal{F} \otimes \mathcal{G}$ such that $(\mathcal{F} \otimes \mathcal{G})_x = \mathcal{F}_x \oplus_{\mathcal{O}_x} \mathcal{G}_x$; one has $\mathcal{O}_x(D + D') = \mathcal{O}_x(D) \otimes \mathcal{O}_x(D')$ for divisors D, D'. The chief interest of sheaf theory is that sheaves of groups may be used to replace the *coefficients* in cohomology groups by "local coefficients" varying with $x \in M$. The cohomology groups $H^j(M, \mathcal{F})$ which one thus defines for each integer $j \geqq 1$ (one also writes $H^j(\mathcal{F})$) have the fundamental property that for any exact sequence of sheaves of abelian groups $0 \to \mathcal{N} \to \mathcal{G} \to \mathcal{G}/\mathcal{N} \to 0$, one has a "long exact sequence"

Theme G

$$(8)\quad 0 \to H^0(\mathcal{N}) \to H^0(\mathcal{G}) \to H^0(\mathcal{G}/\mathcal{N}) \to H^1(\mathcal{N}) \to H^1(\mathcal{G}) \to H^1(\mathcal{G}/\mathcal{N}) \to H^2(\mathcal{N}) \to \cdots$$

Once these new tools were introduced in analysis it was soon recognized that the invariants introduced by the Italian school and by Hodge were easily expressed by sheaf cohomology. In the first place, if M is a compact connected kählerian variety of dimension n, Dolbeault and Serre proved that the corresponding space $H^{r,s}$ of harmonic forms of type (r, s) (see VII-a) is isomorphic to $H^s(\Omega_M^r)$; furthermore, for any divisor D on M, Serre discovered that there is a natural duality pairing the spaces

$$H^j(\mathcal{O}_M(D)) \text{ and } H^{n-j}(\Omega_M^n \otimes \mathcal{O}_M(-D)) = H^{n-j}(\mathcal{O}_M(\Delta - D))$$

"explaining" the intervention of the canonical divisor Δ in Riemann-Roch's theorem (3) (one has written $\Omega_M^n = \mathcal{O}_M(\Delta)$). By definition, the *geometric genus* of M can be written

$$(9)\qquad p_g = \dim(H^0(\Omega_M^n)) \text{ and also } p_g = \dim(H^n(\mathcal{O}_M))$$

by the isomorphism of $H^{r,s}$ and $H^{s,r}$; one has similar invariants for holomorphic exterior forms of all degrees $< n$. The *arithmetic genus* turns out to be the number

$$p_a = \dim H^n(\mathcal{O}_M) - \dim H^{n-1}(\mathcal{O}_M) + \cdots + (-1)^{n-1}\dim H^1(\mathcal{O}_M) \tag{10}$$

and the plurigenera are given by

$$p_k = \dim H^0(\mathcal{O}_M(k\Delta)). \tag{11}$$

In 1937, Eger and Todd introduced, on an algebraic nonsingular projective variety M of complex dimension n, "canonical" equivalence classes of algebraic cycles of dimension $n-j$, which later were recognized to correspond exactly *via* Poincaré duality, to the Chern classes c_j of M; furthermore, Todd discovered that the arithmetic genus of M could be computed by the formula

$$(-1)^n p_a + 1 = \langle T_n(c_1, \cdots, c_n), M \rangle, \tag{12}$$

where T_n is a polynomial with rational coefficients in the Chern classes, defined by the following device: in the power series

$$\prod_{j=1}^{n} \frac{\gamma_j z}{1 - \exp(\gamma_j z)}$$

one considers the coefficient of z^n, which is a symmetric polynomial in the variables γ_j, and one expresses it in terms of the elementary symmetric functions of the γ_j; then one replaces each elementary symmetric function σ_j by c_j. For instance, the first three Todd polynomials are

$$T_1(c_1) = c_1/2, \qquad T_2(c_1, c_2) = (c_2 + c_1^2)/12,$$
$$T_3(c_1, c_2, c_3) = c_2 c_1/24.$$

In 1954, Hirzebruch generalized both Todd's result and the Riemann-Roch formulas of Kodaira by proving that for any divisor D on M, the expression

$$\dim H^0(\mathcal{O}_M(D)) - \dim H^1(\mathcal{O}_M(D)) + \cdots + (-1)^n \dim H^n(\mathcal{O}_M(D))$$

could be expressed as $\langle P(f, c_1, \cdots, c_n), M \rangle$, where f is the first Chern class of the bundle $\mathbf{B}(D)$, and P a polynomial which is obtained by the same device as above, starting from the power series

$$e^{fz} \prod_j \frac{\gamma_j z}{1 - \exp(\gamma_j z)}.$$

It was later recognized that in fact, Hirzebruch's formula was a particular case of a much more general theorem valid for all differential manifolds, the Atiyah-Singer index formula.

The Hirzebruch formula enables one to solve the Riemann-Roch problem when

all cohomology groups $H^j(\mathcal{O}_M(D))$ are reduced to 0 for $j \geqq 1$. Kodaira found sufficient conditions for this fact to hold; for instance, it is true when one replaces D by $D + mH$ where H is the intersection of M and a hyperplane (in the projective space where M is imbedded) and $m > 0$ is large enough. He has also obtained a fundamental criterion for a compact kählerian manifold M to be isomorphic to a projective algebraic variety: there must exist on M a kählerian metric such that the cohomology class of the form Ω (equation (6)) in $H^2(M, \boldsymbol{R})$ belongs to $H^2(M, \boldsymbol{Q})$.

VIII b: The Serre varieties. In 1942, Zariski began a deep study of singularities of projective algebraic varieties over any field, in view of proving a desingularization theorem (which he succeeded to do for dimension $\leqq 3$ and over a field of characteristic 0); for that purpose, he used for the first time the general theory of valuations*, developed 10 years earlier by Krull. In the course of this work, he introduced the generalization of the "abstract Riemann surface" of Dedekind-Weber for an arbitrary field K of algebraic functions over a field k, defining it to be the set V of all valuations of K which vanish on k^*; but in addition, using ideas introduced a few years earlier by M. Stone, he defined on V (by purely algebraic considerations) a *topology* for which V became quasi-compact, although that topology is not Hausdorff in general: for instance, in the case of dimension 1, considered by Dedekind-Weber, the closed sets are V and all the finite subsets of V.

Theme C

By 1950 A. Weil observed that this "Zariski topology" could be defined on his "abstract varieties" (see VII-b); not only did it appreciably improve the exposition of the theory by allowing one to use a "geometric" language, but it also made possible a definition of *vector bundles* modeled on the classical one, and to extend to abstract varieties the relations between divisors and line bundles (see VIII-a). Going one step further, Serre, in 1955, had the idea to transfer in the same way the theory of sheaves to abstract varieties, using the Zariski topology instead of the usual one in Leray's definition. At the same time, he observed that the concept of sheaf made possible a much simpler definition of "abstract varieties," using the general idea of "ringed space" of H. Cartan, i.e., a topological space X on which is given a sheaf of rings $\mathcal{O}_X$; the advantage of this kind of structure is that it lends itself very easily to "gluing" ringed spaces along open subsets, the verification of the conditions of compatibility being usually trivial. In Serre's case the "pieces" which are glued together are *affine varieties* over an algebraically closed field k of

Theme G

---

* The only difference between the definition of a general valuation and the definition of a discrete valuation (see VI–a) is that the valuation may take its value in an *arbitrary* totally ordered group. For instance, the group $\boldsymbol{Z} \times \boldsymbol{Z}$ may be totally ordered by writing $(m, n) < (m', n')$ if either $m < m'$, or $m=m'$, and $n<n'$ ("lexicographic ordering"); one may then define on $\boldsymbol{C}(X, Y)$ a valuation with value in that totally ordered group by taking for $w(P)$, where P is a polynomial $\neq 0$, the smallest (m, n) in $\boldsymbol{Z} \times \boldsymbol{Z}$ for which the term in $X^m Y^n$ in P has a nonzero coefficient.

arbitrary characteristic: such a variety X is a (Zariski) closed set of some k^n (i.e., defined by polynomial equations), and $\mathcal{O}_X$ is the sheaf of rings such that for each open set $U \subset X$, $\mathcal{O}(U) = H^0(U, \mathcal{O}_X)$ consists of the restrictions to U of the rational functions $P(X)/Q(X)$ on k^n which are defined (i.e., $Q(x) \neq 0$) at every $x \in U$. Of course cohomology groups $H^j(\mathcal{F})$ can still be defined when $\mathcal{F}$ is a sheaf of modules over the rings $\mathcal{O}_x$; they are vector spaces over k and Serre computed the groups $H^j(\mathcal{O}_M(mH))$ for $M = \boldsymbol{P}_n(k)$ and H a hyperplane ($m \in \boldsymbol{Z}$); he also extended to arbitrary fields and to projective varieties his duality theorem; but when k has characteristic $p > 0$, most of the results obtained in the classical case by the methods of Lefschetz and Hodge fail to generalize: for instance, the dimension of $H^r(\Omega_X^s)$ and of $H^s(\Omega_X^r)'$ for a projective variety X are not necessarily equal. Nevertheless, Grothendieck and Washnitzer were able independently to extend Hirzebruch's formula to fields k of arbitrary characteristic, and Grothendieck, by the introduction of his "K-theory," gave a far reaching generalization of that formula. Finally, when k is the complex field, Serre showed that the cohomology groups obtained by using the Zariski topology coincided with the classical ones.

Being chiefly interested in cohomology, Serre did not dwell at length on the general properties of his varieties; these were investigated in detail by Chevalley almost simultaneously (in a different language, which we do not reproduce here). One of the points which should be emphasized is that with Serre and still more with Chevalley, birational geometry fades out of the picture and the concept of *morphism* comes to the fore. Until then, the center of interest was the theory of *complete* varieties, and it is only seldom that a correspondence between two such varieties X, Y, even if it assigns only one point of Y to a point of X (a $(1, n)$-correspondence in classical language), is defined at *every* point of X. A morphism $f: X \to Y$, where X and Y are Serre varieties, is on the contrary a *mapping of* X into Y, which is continuous for the Zariski topologies and such that for every point $x \in X$ and every affine neighborhood V of $y = f(x)$, there is an affine neighborhood U of x such that $f(U) \subset V$ and, for every function $s \in H^0(V, \mathcal{O}_Y)$, the function $x \mapsto s(f(x))$ defined in U, belongs to $H^0(U, \mathcal{O}_X)$. The main results of Chevalley are general theorems on morphisms and studies of special types of morphisms using results of commutative algebra going back to E. Noether and Krull. It had been known for a long time that the image $f(X)$ of X by a morphism $f: X \to Y$ was not even locally closed in Y in general; Chevalley showed however that when X is irreducible, $f(X)$ always contains a set which is open and dense in the subspace $\overline{f(X)}$ of Y. Another of Chevalley's results is that if X and Y are irreducible, and for each $x \in X$ one writes $e(x)$ the maximum of the dimensions of the irreducible components of $f^{-1}(f(x))$ which contain x, then the mapping $x \mapsto e(x)$ is upper semi-continuous in X (in other words, when x' is close enough to x, $e(x')$ is never $< e(x)$).

Theme B

Chevalley also showed how important concepts introduced by Zariski in the 1940's, and which A. Weil had already used in his theory of abstract varieties, led to

very suggestive theorems on morphisms. For projective varieties, Zariski had observed that the "regularity" properties of a point $x \in X$ were linked very closely to the structure of the *local ring* $\mathcal{O}_x$ of the variety X at that point: x only belongs to one irreducible component if $\mathcal{O}_x$ has no zero divisors, and x is *simple* if $\mathcal{O}_x$ is a *regular* local ring (i.e., $\mathcal{O}_x$ is an integral domain whose field of fractions has a transcendence degree over the base field k (always assumed to be algebraically closed) equal to the dimension over k of the vector space $\mathfrak{m}_x/\mathfrak{m}_x^2$, where $\mathfrak{m}_x$ is the maximal ideal of $\mathcal{O}_x$). A property, of which Zariski was the first to grasp the geometric significance, is the fact for $\mathcal{O}_x$ to be *integrally closed* in its field of fractions, in which case x is said to be *normal.* Zariski showed that simple (or normal) points of an irreducible variety formed an open dense set, and that the complement of the set of normal points has codimension *at least* 2. Furthermore, Zariski defined for each projective irreducible variety X its "normalization;" this can easily be extended to Serre varieties: for any finite extension L of the field of rational functions K of X, there is a variety X' and a morphism p: $X' \to X$ such that for each affine open set U of X, $p^{-1}(U)$ is an affine open set of X' and the ring $H^0(p^{-1}(U), \mathcal{O}_{X'})$ is the integral closure in L of the ring $H^0(U, \mathcal{O}_X)$; X' is called the normalization of X *in L,* and simply the normalization of X if $L = K$. The normalization of X is of course birationally equivalent to X, and its singular points form a subvariety of codimension $\geqq 2$; in particular, if X is a curve, X' has no singular points, and this is the simplest "desingularization" of a curve (valid in every characteristic).

The climax of Zariski's investigations on normal varieties had been his "main theorem" expressed in the language of birational correspondences; Chevalley showed that it implies a far more intuitive result about morphisms: suppose X and Y are irreducible and *normal* varieties, f: $X \to Y$ is a morphism such that $f(X)$ is dense in Y and each set $f^{-1}(y)$ is *finite* for $y \in Y$. Then f factorizes in $X \xrightarrow{g} Y' \xrightarrow{p} Y$ where Y' is the normalization of Y in the field of rational functions of X, and g is an *isomorphism* of X onto an *open* subvariety of Y'.

Finally, Chevalley defined the notion of *complete* variety in a much simpler way than before: X is complete if, for every variety Y, the second projection $X \times Y \to Y$ is a *closed* mapping.

The interest of Chevalley in such theorems was spurred by the theory of *algebraic groups*, which he and A. Borel brought to a high level of development during the 1950's; in that theory, both affine and complete varieties play an important part and the preceding theorems are powerful tools.

VIII c: Schemes and topologies. Until the 1950's, no one seems to have tried to give an *intrinsic* definition of an affine variety over an *algebraically closed field k*, independent of any imbedding of the variety in some "affine space" k^n, although the tools to do so were available since the 1890's. In his work on invariant theory, Hilbert had proved his famous "*Nullstellensatz*," one of the forms of which is that the maximal ideals of the algebra of polynomials $k[X_1, \cdots, X_n]$ are in one-to-

one correspondence with the elements $z = (\zeta_1, \cdots, \zeta_n) \in k^n$, such an element corresponding to the ideal generated by the polynomials $X_1 - \zeta_1, \cdots, X_n - \zeta_n$. Just as Riemann attached to a projective curve the field of rational functions on that curve, so one may attach to an affine variety $V \subset k^n$ the ring $R(V)$ of the restrictions to V of all *polynomial* functions on k^n; this ring is a finitely generated algebra over k, which has no nilpotent elements (one says it is *reduced*); and by Hilbert's *Nullstellensatz*, the points of V are in one-to-one correspondence with the maximal ideals of $R(V)$. Conversely, it is readily seen that *any* reduced and finitely generated k-algebra has the form $R(V)$ for an affine variety determined up to isomorphism. Furthermore, when V is irreducible, it is even possible to define the sheaf $\mathcal{O}_V$ directly from the ring $R(V)$: for any open (Zariski) subset U of V which is defined as the set of points x such that $f(x) \neq 0$ for some $f \in R(V)$, one defines $\mathcal{O}(U)$ as the ring of rational functions of type g/f^m for $g \in R(V)$ and m a positive integer, and it is easy to see that this defines completely $\mathcal{O}_X$. Finally, if V, W are two affine varieties over k, we have seen above that to a morphism $f\colon V \to W$ corresponds a k-algebra homomorphism $R(f)\colon R(W) \to R(V)$; but the converse is also true, for Hilbert's *Nullstellensatz* implies that for any such homomorphism $\phi\colon R(W) \to R(V)$, the inverse image $\phi^{-1}(\mathfrak{m})$ of a maximal ideal of $R(V)$ is again a maximal ideal in $R(W)$, and $\mathfrak{m} \mapsto \phi^{-1}(\mathfrak{m})$ is the morphism corresponding to ϕ. In the language of categories, which was beginning to be used in the late 1950's, the category of affine varieties over k was *equivalent* to the *dual* of the category of reduced finitely generated (commutative) k-algebras.

Following a suggestion of Cartier, A. Grothendieck undertook around 1957 a gigantic program aiming at a vast generalization of algebraic geometry, absorbing all previous developments and starting from the category of *all* commutative rings (with unit) instead of reduced finitely generated algebras over an algebraically closed field. If one wanted to define a category which would be equivalent to the dual of the category of all commutative rings, a nontrivial modification was needed from the start, since if $\phi\colon A \to B$ is a homomorphism of rings (sending unit element on unit element), the inverse image $\phi^{-1}(\mathfrak{m})$ of a maximal ideal of B is not in general a maximal ideal of A, whereas the inverse image $\phi^{-1}(\mathfrak{P})$ of a *prime* ideal of B is always a prime ideal of A. It was thus necessary to take as the set replacing the affine variety the *spectrum* of A, i.e., the set $\mathrm{Spec}(A)$ of all *prime* ideals of A; closed sets in $\mathrm{Spec}(A)$ are defined as sets of prime ideals containing a given (arbitrary) ideal of A, hence a "Zariski topology" for which, however, finite sets are no longer closed in general; finally, using work of Chevalley and Uzkov on localization dating from the 1940's, it is possible to give a meaning to g/f^m even when f is a zero-divisor of A, hence to define the sheaf $\mathcal{O}_X$ on $X = \mathrm{Spec}(A)$ in the same way as for affine varieties. The ringed spaces thus obtained are called *affine schemes* and they form a category equivalent to the dual of the category of all commutative rings; finally, the usual "gluing process" for ringed spaces yields the category of *schemes* by replacing affine varieties by affine schemes.

The experience of the last 10 years has convinced the specialists that, in spite of the much greater amount of commutative algebra techniques which it requires, the theory of schemes is the context in which the problems of algebraic geometry are best understood and attacked. Among the features which distinguish it from previous conceptual frames for algebraic geometry, let us mention only the few following ones:

(1) The notion of *generic point*, which had disappeared from the Serre-Chevalley theory, is now reintroduced in a natural way: for instance, if A is an integral domain, its (unique) generic point is the prime ideal (0) in $\mathrm{Spec}(A)$; its "generic" property is expressed by the fact that its *closure* is the *whole* space $\mathrm{Spec}(A)$, and thus continuity arguments in the Italian style (but in the Zariski topology!) are now again available.

(2) The predominance of "relative" versus "absolute" notions, or, put in a different way, the fact that most of the times what is studied is not a scheme but a *morphism* of schemes $f: X \to S$, where S is often quite arbitrary (one also says that the study of such morphisms, for fixed S, is the study of "S-schemes"). This is particularly apparent when it comes to imposing *finiteness conditions* (without *any* such condition, there is very little likelihood of ever getting any deep result): Grothendieck has shown that, except for cohomological notions, one may usually allow the "base scheme" S to be free from finiteness assumptions (such as being noetherian, or of finite dimension, etc.), and the results only depend on finiteness conditions for the morphism f; this allows considerable freedom in the "change of bases" (see below).

(3) Given two "S-schemes" $f: X \to S$, $g: Y \to S$, there is an essentially unique triplet consisting in an S-scheme $X \times_S Y$ and two morphisms $p_1: X \times_S Y \to X$, $p_2: X \times_S Y \to Y$ such that $f \circ p_1 = g \circ p_2$, which is the "categorical" *product* of X and Y over S: this means that, given two morphisms $u: Z \to X$, $v: Z \to Y$ such that $f \circ u = g \circ v$, there is a unique morphism $w: Z \to X \times_S Y$ such that $u = p_1 \circ w$ and $v = p_2 \circ w$ (there is no similar result for Serre varieties; it easily follows from the existence of the tensor product $B \otimes_A C$ of arbitrary A-algebras, where A is any ring).

Most of the time this fundamental process is applied to study the morphism $f: X \to S$ by replacing the "base" S by another one Y, in such a way that the new morphism p_2, which is now written $f_{(Y)}: X_{(Y)} \to Y$ (the notation $X_{(Y)}$ replacing $X \times_S Y$) can be more easily handled. This "change of base" is probably the most powerful tool in the theory of schemes, generalizing in a bewildering variety of ways the old idea of "extending the scalars." To give only one example, consider at any point $s \in S$ the residual field $k(s) = \mathcal{O}_s/\mathfrak{m}_s$ of the local ring $\mathcal{O}_s$ at that point; then $X_s = X \times_S \mathrm{Spec}(k(s))$ has as underlying space the "fiber" $f^{-1}(s)$ in X and (provided f satisfies finiteness conditions) it can be considered as a "variety" over the field $k(s)$ (in a slightly more general sense than with Serre). In this way, an S-scheme X may be considered as a "family of varieties" X_s parametrized by S

(generalizing the old Picard method (see VI-c)) and many properties of S-schemes may be obtained by a study of the fibers X_s.

(4) It may seem strange at first that one should consider affine schemes $\mathrm{Spec}(A)$ even when A has *nilpotent elements* other than 0; but in fact, this also corresponds to geometric facts which were not taken into account by older theories. For instance, consider the parabola $y^2 - x = 0$ in $\boldsymbol{C}^2$ and the mapping which projects it on the x-axis; in the language of schemes, we consider the affine schemes $U = \mathrm{Spec}(\boldsymbol{C}[X, Y]/(Y^2 - X))$, $V = \mathrm{Spec}(\boldsymbol{C}[X])$ and the morphism $p\colon U \to V$ which corresponds to the natural injection $\boldsymbol{C}[X] \to \boldsymbol{C}[X, Y]/(Y^2 - X)$ which sends X onto the class of X. A maximal ideal $(X - \zeta)$ in $\boldsymbol{C}[X]$ is identified with the point $\zeta \in \boldsymbol{C}$, and the fiber $V_\zeta = p^{-1}(\zeta)$ is the affine scheme $\mathrm{Spec}(\boldsymbol{C}[Y]/(Y^2 - \zeta))$; now, if $\zeta \neq 0$, the ring $\boldsymbol{C}[Y]/(Y^2 - \zeta)$ is isomorphic to the direct sum of two fields isomorphic to $\boldsymbol{C}$, corresponding to the fact that the fiber has two distinct points; but if $\zeta = 0$, $\boldsymbol{C}[Y]/(Y^2)$ has nilpotent elements: the two points have become "infinitely near" one another. It turns out that this is a general phenomenon: nilpotent elements in the local rings of a scheme are the algebraic counterpart of "infinitesimal" properties, and their presence allows a much more natural and flexible treatment of these properties than in classical algebraic geometry (see e.g. [**8**]).

(5) If we return to the concept of affine Serre variety, corresponding to a reduced finitely generated algebra A over an algebraically closed field k, the points of the variety are not *all* points of $\mathrm{Spec}(A)$, but only the *closed* ones, corresponding to all homomorphisms $A \to k$ which are k-homomorphisms, i.e., such that the composition with the natural mapping $k \to A$ gives the identity on k; similarly, if one wants to consider the points of variety "with coordinates in a field K extension of k" (see VII-b), one has to consider homomorphisms $A \to K$ which by composition $k \to A \to K$ give the homomorphism defining the extension K of k. This idea has been greatly generalized by Grothendieck: for an S-scheme $X \to S$ the "points of X in an arbitrary S-scheme T" (or more briefly the "T-points" of X) are by definition the morphisms $T \to X$ which, composed with $X \to S$, give the structural morphism $T \to S$; if we denote by $\mathrm{Mor}_S(T, X)$ the set of these "S-morphisms," it can easily be shown that $T \mapsto \mathrm{Mor}_S(T, X)$ is a *functor* from the category of S-schemes to the category of sets, and that the knowledge of that functor entirely determines the S-scheme X, which is said to "represent" the functor. This idea has become a very fruitful principle allowing the definition of schemes by the functor which they "represent," which is generally much easier (provided one has general theorems establishing the "representability" of functors); in particular, one transfers in that way to the theory of schemes many classical constructions such as projective spaces, Grassmannians, Chow varieties, Picard varieties, and one is able to give a general meaning to the concept of "moduli" introduced by Riemann for curves.

(6) It was early recognized that the Zariski topology on schemes had some unpleasant features regarding "vector bundles:" natural definitions of S-schemes $X \to S$, which in classical geometry gave vector bundles X over S, did not have in

general the property of being "locally" products of a (Zariski) neighborhood and a "typical fiber" (one says that they are not "locally trivial" for the Zariski topology). However, Serre observed that in important cases, a mild "extension of the base" $T \to S$, where T is an "etale covering" of S (which corresponds in classical geometry to an unramified covering with finitely many sheets) was enough to restore "local triviality." Starting from this remark, Grothendieck conceived the idea of replacing the Zariski topology on S by a new structure, called "*etale topology*," which is not any more a topology in the usual sense; essentially it consists in replacing the usual open subsets of S (or rather their natural injections $U \to S$) by etale coverings of S (one may say that the open sets are now "out of the space" instead of being parts of it). The important fact is that he was able to transfer to this new concept the definition of sheaves and of sheaf cohomology, and to show that this "etale cohomology" can partly remedy to the defects of the usual (Zariski) sheaf cohomology for varieties over a field of characteristic $p > 0$.

IX. OPEN PROBLEMS

To have some idea of the dozens of problems on which algebraic geometers are now working, one may consult for instance the various reports in [**18**], [**19**], or [**20**]. We will conclude by mentioning very briefly some of the most conspicuous ones.

(1) The famous problem of "desingularization" of algebraic varieties over a field k has been solved by Hironaka in all dimensions, when k has characteristic 0, and this result has become a very powerful tool in many problems of algebraic geometry, both classical and "abstract." For fields of characteristic $p > 0$, the problem is still open in dimensions $\geqq 3$; for dimension 2, the desingularization theorem has been proved by Abhyankar in all characteristics.

(2) The problem of Riemann's "moduli" has attracted much attention during the last 20 years, both in classical and in abstract geometry: the general idea is to prove the existence of a variety (or scheme) whose points would correspond to isomorphism classes of curves of a given genus over a given field; the most comprehensive results to date are those of Mumford, who has proved the existence of such a scheme; but much remains to be done regarding the properties of that scheme. One has similar results when curves of given genus are replaced by abelian varieties of given dimension; but already for algebraic surfaces, very little progress has been made on similar problems. Even when one considers "local" problems, i.e., how algebraic structures depending on parameters may "deform" in the neighborhood of a point in the parameter space, the results are far from final.

(3) In spite of the progresses brought by "etale cohomology" (and other similar theories based on other types of "Grothendieck topologies"), the cohomological properties of varieties over a field of characteristic $p > 0$ are not yet well understood, and nothing has yet satisfactorily replaced the abelian integrals in that case. Central in these problems are the "Weil conjectures" which he formulated as extensions to algebraic varieties of arbitrary dimension of his work on the zeta function of algebraic curves over finite fields; some of them have been proved by Grothendieck

and M. Artin, using etale cohomology, but the extension of the "Riemann hypothesis" has up to now resisted all efforts.

(4) In classical algebraic geometry, the theory of integrals of "second" of "third" kinds on projective algebraic varieties of arbitrary dimension is still incomplete, although much advanced recently by the work of Leray, Hodge-Atiyah and Griffiths on the concept of "residue." Generalizations of the Hodge theory to non compact algebraic varieties (over $\boldsymbol{C}$) with singularities have recently been started by Deligne and others.

(5) One would expect that the precise knowledge of divisors under various "equivalence" concepts (see VII-d) should extend to "cycles" of arbitrary condimension, but even in the classical case that theory is still in an embryonic stage.

(6) Finally, the beautiful results of Castelnuovo and Enriques on the characterization of classes of surfaces by properties of their invariants have been greatly extended by Kodaira and Shafarevich [**11**], and generalized by Mumford to surfaces over an algebraically closed field of characteristic $p > 0$ [**19**], but much remains to be done, and practically no comparable results have been obtained in higher dimensions.

References

1. M. Baldassari, Algebraic varieties, Ergeb. der Math., Heft 12, Springer, Berlin-Göttingen-Heidelberg, 1956.

2. J. Dieudonné, Algebraic geometry, Advances in Math., 3 (1969) 233–321.

3. ———, Fondements de la géométrie algébrique moderne, Advances in Math., 3 (1969) 322–413.

4. F. Enriques, Sui sistemi continui di curve appartenenti ad una superficie algèbrica, Comm. Math. Helv., 15 (1943) 227–237.

5. F. Hirzebruch, Topological methods in algebraic geometry, Springer, Berlin-Heidelberg-New York, 3rd ed., 1966.

6. S. Lefschetz, L'Analysis Situs et la Géométrie algébrique, Gauthier-Villars, Paris, 1924.

7. D. Mumford, Geometric invariant theory, Erg. der Math. Heft 34, Springer, Berlin-Heidelberg-New York, 1965.

8. ———, Lectures on curves on an algebraic surface, Princeton Univ. Press, Princeton, 1966.

9. ———, Abelian varieties, Oxford Univ. Press, Oxford, 1970.

10. F. Severi, Intorno ai sistemi continui di curve sopra una superficie algèbrica, Comm. Math. Helv., 15 (1943) 238–248.

11. I. Shafarevich et al. Algebraic surfaces, Proc. Steklov Inst. of Math., Amer. Math. Soc., 1967.

12. B. L. van der Waerden, Einführung in die algebraische Geometrie, Springer, Berlin, 1939.

13. A. Weil, Foundations of algebraic geometry, Amer. Math. Soc. Coll. Publ., 29 (1946).

14. ———, Sur les courbes algébriques et les variétés qui s'en déduisent, Hermann, Paris, 1948.

15. ———, Introduction à l'étude des variétés kählériennes, Hermann, Paris, 1958.

16. O. Zariski, Algebraic surfaces, Erg. der Math. 2nd ed., Springer, Berlin-Heidelberg-New York, 1971.

17. ———, An introduction to the theory of algebraic surfaces, Lecture Notes in Math., 83, Springer, Berlin-Heidelberg-New York, 1969.

18. *Dix exposés sur la cohomologie des schémas*, North-Holland, Amsterdam-London, 1968.

19. *Global Analysis* (*Papers in honor of K. Kodaira*), Princeton Univ. Press, 1969.

20. *Actes du Congrès international des mathématiciens*, Nice, 1970, vol. I et II, Gauthier-Villars, Paris, 1971.

CRUDELY STATIONARY COUNTING PROCESSES

KAI LAI CHUNG, Stanford University

1. Introduction. The theorems by Khintchine, Korolyuk, and Dobrushin in the theory of stationary point processes are basic and simple theorems. Korolyuk's theorem was originally derived from the Palm-Khintchine formulas; a direct proof was given in Cramér-Leadbetter [1]. Its real simplicity seems to be obscured by the slightly complicated presentation of the proof. The same may be said of the proof of Dobrushin's theorem involving an unnecessary contraposition as well as some epsilonics. Both results become quite transparent when dealt with by standard methods of measure and integration in sample space. After all, these are problems of probability theory and nowadays students spend a lot of time learning this kind of "abstract" set-up. It would be a pity not to use the knowledge so acquired in straightforward situations such as these theorems. In doing so we arrive at certain natural extensions which seem to put the results in proper perspective. The results in R^d, obtained by the same method, seem to be new.

The reader is referred to Leadbetter [3] for another simple approach, which came belatedly to our attention.

2. Definitions and statements. Let $(\Omega, \mathscr{F}, P)$ be a probability space. Each ω in Ω is a set $S(\omega)$ of points on $R = (-\infty, +\infty)$ endowed with "multiplicity," namely a positive integer attached to the point. A point with **multiplicity** m is counted as m ordinary points at distance zero to each other; it will be called a **multiple point** when $m \geqq 2$. The fundamental assumptions are as follows:

(A) For each finite interval I in R, the "number" of points in $S(\omega) \cap I$, counted with their multiplicities, is finite. This number will be denoted by $N(I, \omega)$; as usual $N(I)$ is the function $\omega \to N(I, \omega)$.

(B) The function

$$(s, t, \omega) \to N([s, t]; \omega),$$

where $s \leqq t$ is measurable with respect to the product field $\mathscr{B} \times \mathscr{B} \times \mathscr{F}$ where $\mathscr{B}$ is the Euclidean Borel field on R.

It follows that for each interval I, $N(I)$ is a random variable. We do not define $N(\cdot)$ for other sets than intervals.

The collection $\{N(I, \omega)\}$ with I ranging over intervals and ω over Ω, will be called a **counting process** on R. It is said to be **crudely stationary** iff whenever I and J

Since receiving his Princeton Doctor's Degree, Kai Lai Chung has held positions at Princeton, Cornell, Columbia, Syracuse, Chicago, University of Illinois, and Stanford. He has spent leaves abroad at the University of Strasbourg and the ETH Zurich. His main research interest is probability and he is the author of *Markov Chains*, *Course in Probability Theory*, and *Boundary Theory*. *Editor.*

are two compact intervals of equal length the random variables $N(I)$ and $N(J)$ have the same distribution. The same will then be true for any two finite intervals of equal length, whether closed or open or half-open and possibly degenerate, by Proposition 1 below. Indeed the sole purpose of the formulation above is to bring that proposition into question. The adjective "crude" is used to distinguish it from "strict," which requires much more (see [**1**], 3.8).

An equivalent formulation is to define an integer-valued stochastic process $\{X(t,\omega);\, t\in R,\ \omega\in\Omega\}$ as follows:

$$X(t,\omega) = \begin{cases} N([0,\,t),\omega) & \text{if} \quad t \geqq 0, \\ -N([t,\,0),\omega) & \text{if} \quad t < 0. \end{cases}$$

For each ω, $t \to X(t,\omega)$ is then a right continuous purely jumping non-decreasing function. The set of its jump-points is $S(\omega)$ and the size of jump at each point is its multiplicity. If X has strictly stationary increments in the usual sense, then the **increment process** $\{N(I,\omega)\}$ will be not only crudely, but even strictly stationary. While the conversion to X has the advantage of making a counting process into a more standard object, the language and notation for N is slightly more direct and so preferred here. We begin by settling a small point, which in the strictly stationary case follows from the fact that the Borel-Lebesgue measure is the unique translation-invariant measure on R, apart from a constant factor.

We use E below to denote the mathematical expectation, and write, e.g., $\{N([t,t])=0\}$ for $\{\omega \mid N([t,t],\,\omega)=0\}$.

PROPOSITION 1. *For each degenerate interval* $[t,\,t]$ *we have*

$$P\{N([t,t])=0\}=1. \tag{1}$$

Proof. The set

$$H=\{(t,\omega)\,|\,N([t,t],\omega)\neq 0\}=\{(t,\omega)\,|\,t\in S(\omega)\}$$

belongs to $\mathscr{B}\times\mathscr{F}$ by (B). Integrating its indicator 1_H over $[0,1]\times\Omega$ and applying Fubini's theorem, we obtain in view of (A):

$$0=\int_\Omega 0\,P(d\omega)=\int_\Omega\left[\int_{[0,1]}1_H(t,\omega)\,dt\right]P(d\omega)$$
$$=\int_{[0,1]}\left[\int_\Omega 1_H(t,\omega)\,P(d\omega)\right]dt=\int_{[0,1]}E\{N([t,t])\}\,dt.$$

By crude stationarity, $t\to E\{N([t,t])\}$ is a constant c, where $0\leqq c\leqq+\infty$; hence $c=0$ which is equivalent to (1).

PROPOSITION 2. *Either* (i) $E\{N(I)\}=\infty$ *for every non-degenerate* I; *or* (ii)

$E\{N(I)\} < \infty$ for every finite I. In case (ii), *for any sequence $\{I_n\}$ such that $I_n \downarrow$ and $|I_n| \to 0$ (where $|I|$ = length of I), we have*

$$P\{\lim_n N(I_n) = 0\} = 1. \tag{2}$$

Furthermore, in either case we have for every $t \geqq 0$:

$$E\{N([0,t])\} = E\{N([0,1])\}\, t \tag{3}$$

provided we set $\infty \cdot 0 = 0$.

Proof. Observe that (2) is false in case (i) even when $I_n \downarrow \varnothing$. The rest follows from crude stationarity, dominated convergence and Proposition 1, and we omit the details of a familiar argument.

From now on we shall write

$$N(t) = N([0,t]),\ \mu(t) = E\{N(t)\},\ \mu = \mu(1);$$

so that (3) becomes

$$\mu(t) = \mu t \text{ where } 0 \leqq \mu \leqq \infty. \tag{4}$$

Furthermore we introduce the notation for $k \geqq 1$:

$$\begin{aligned} p_k(t) &= P\{N(t) = k\}, \\ r_k(t) &= P\{N(t) \geqq k\} = \sum_{j=k}^{\infty} p_j(t), \\ \lambda_k &= \lim_{t \downarrow 0} \frac{r_k(t)}{t}, \end{aligned}$$

whenever the limit exists. The process is said to be **regular** when $\lambda_2 = 0$.

The theorems by Khintchine, Dobrushin, and Korolyuk may be stated as follows (originally given for the strictly stationary case).

KHINTCHINE'S THEOREM. *λ_1 always exists: $0 \leqq \lambda_1 \leqq \infty$.*

DOBRUSHIN'S THEOREM. *If $\mu < \infty$ and there are no multiple points, then $\lambda_2 = 0$.*

KOROLYUK'S THEOREM. *If $\lambda_2 = 0$, then $\lambda_1 = \mu \leqq \infty$.*

It is the object of this note to formulate natural extensions of these results and give very simple proofs of them.

PROPOSITION 3. *If for some $k \geqq 1$ we have $\lambda_{k+1} = 0$, then*

$$\mu = \lim_{t \downarrow 0} \left(\sum_{j=1}^{k} \frac{r_j(t)}{t} \right). \tag{5}$$

For $k=1$ this reduces to Korolyuk's theorem, which contains Khintchine's in the regular case. In general, the existence of λ_j for $2 \leqq j \leqq k$ is neither postulated nor implied. It is known (Khintchine [2], 3.8) that all λ_k exist for a strictly stationary process *without after effect*, that is, a **compound Poisson process** (see below).

PROPOSITION 4. *If all λ_k exist for $2 \leqq k < \infty$, finite or infinite, then*

$$\mu = \sum_{k=1}^{\infty} \lambda_k. \tag{6}$$

PROPOSITION 5. *Let $k \geqq 1$. If $\mu < \infty$ and there are no points with multiplicity $\geqq k+1$, then $\lambda_{k+1} = 0$. The converse is true for $\mu \leqq \infty$.*

For $k=1$, the first part is Dobrushin's theorem; the second part is trivial for any k (cf. Cramér-Leadbetter [1], page 54).

PROPOSITION 6. *There is a strictly stationary counting process without any multiple point for which*

$$\mu = \lambda_1 = \infty, \quad 0 < \lambda_k < \infty \textit{ for } k \geqq 2. \tag{7}$$

Further relevant facts will be mentioned at the end of section 3.

3. Proofs of the propositions. Let us begin by writing the elementary formula

$$E\{N(t)\} = \sum_{k=1}^{\infty} P\{N(t) \geqq k\}$$

in terms of our notation above:

$$\mu = \frac{\mu(t)}{t} = \sum_{k=1}^{\infty} \frac{r_k(t)}{t}. \tag{8}$$

We may thus regard the announced propositions as a study of the limiting form of the relation (8) as we let $t \downarrow 0$ and try to take the limits inside the summation—a meet game in analysis, made interesting here by the probabilistic interpretations.

Proof of Proposition 3. For each $t > 0$, we have (with an obvious abridging of notation)

$$N[0,1-t] \leqq \sum_{n=0}^{m-1} N[nt,(n+1)t], \qquad m = \left[\frac{1}{t}\right].$$

It is plain that for each integer $M > 0, \{N[0,1] \leqq M\} \subset \{N[nt,(n+1)t] \leqq M\}$ for $0 \leqq n \leqq m-1$. Hence we have

$$\begin{aligned}\int_{N[0,1] \leqq M} N[0,1-t]\, dP &\leqq \sum_{n=0}^{m-1} \int_{N[nt,(n+1)t] \leqq M} N[nt,(n+1)t]\, dP \\ &= m \int_{N[0,t] \leqq M} N[0,t]\, dP\end{aligned}$$

by crude stationarity in the last equation. From the definitions of the quantities involved, we have

$$\int_{N[0,t]\leqq M} N[0,t]\,dP = \sum_{j=1}^{M} j\,p_j(t) \leqq \sum_{j=1}^{M} r_j(t).$$

It follows that

$$\int_{N[0,1]\leqq M} N[0,1-t]\,dP \leqq \frac{t+1}{t} \sum_{j=1}^{M} \frac{r_j(t)}{t}. \tag{9}$$

Letting $t\downarrow 0$ and using the monotone convergence theorem on the left together with Proposition 1, we obtain

$$\int_{N[0,1]\leqq M} N[0,1]\,dP \leqq \varliminf_{t\downarrow 0} \sum_{j=1}^{k} \frac{r_j(t)}{t} \tag{10}$$

because the hypothesis $\lambda_{k+1} = 0$ forces $\underline{\lambda_j} = 0$ for $j \geqq k+1$. Letting $M\uparrow\infty$ and observing that the reverse inequality for $\overline{\lim}_{t\downarrow 0}$ is trivial we get (5).

Proof of Proposition 4. It is plain from (8) (a case of Fatou's lemma) that

$$\mu \geqq \sum_{k=1}^{\infty} \lambda_k. \tag{11}$$

Letting $t\downarrow 0$ in (9) as before, then $M\uparrow\infty$, we obtain the reverse of (11) and so (6).

Proof of Proposition 5. Fix k and define first for each interval I:

$$\xi(I) = 1_{[N(I)\geqq k+1]}$$

and then for $t > 0$:

$$\eta(t) = \sum_{n=0}^{m-1} \xi[nt,(n+1)t], \qquad m = \left[\frac{1}{t}\right].$$

Thus $\eta(t)$ is the number of subintervals $[nt,(n+1)t]$ in which there are at least $k+1$ points counted with multiplicity. If no point has multiplicity $\geqq k+1$, then each $S(\omega)$ is a discrete set of "points with multiplicity $\leqq k$." If $\delta(\omega)$ denotes the minimum distance between the points of $S(\omega)\cap I$ without their multiplicities, then $\eta(t,\omega) = 0$ for $0 < t < \delta(\omega)$. Thus

$$P\{\lim_{t\downarrow 0} \eta(t) = 0\} = 1. \tag{12}$$

On the other hand, it is obvious that $\eta(t) \leqq N([0,1])$, where the right member above has expectation μ. If $\mu < \infty$ then by dominated convergence

$$\lim_{t\downarrow 0} E\{\eta(t)\} = 0.$$

Now we have by crude stationarity

(13) $$E\{\eta(t)\} = \left[\frac{1}{t}\right] E\{\xi[0,t]\} = \left[\frac{1}{t}\right] r_{k+1}(t).$$

Hence as $t \downarrow 0$ the last term tends to 0, i.e., $\lambda_{k+1} = 0$.

The converse will be shown in an extended form in Proposition 8 below.

Normally speaking, the condition $\lambda_k > 0$ should signal the existence of points with multiplicity $\geqq k$. This is the case for a compound Poisson process, which may be derived from a simple Poisson process by randomizing the multiplicity of each point according to a fixed distribution and independently over all the points. It is also the case, in a more general way, for a continuous time homogeneous Markov chain, where the situation is indicated by the formula, in standard notation:

$$\lim_{t \downarrow 0} \frac{p_{jk}(t)}{t} = q_{jk}, \qquad j \neq k.$$

Proposition 5 says that it is always true for a crudely stationary counting process when $\mu < \infty$. It may or may not be surprising that this is no longer so when $\mu = \infty$, as shown in the following counterexample.

Example (*proof of Proposition* 6). For each $\lambda > 0$ let $N^{(\lambda)}$ be a (simple) Poisson process on R with *intensity* $\lambda_1 = \lambda$; namely,

$$\begin{aligned} \mu^{(\lambda)}(t) &= E\{N^{(\lambda)}[0,t]\} = \lambda t; \\ r_{k+1}^{(\lambda)}(t) &= 1 - e^{-\lambda t} \sum_{j=0}^{k} \frac{(\lambda t)^j}{j!}, \qquad k \geqq 0. \end{aligned}$$

Let N denote the counting process obtained by randomizing λ according to the distribution F. Specifically, we choose F to have the density f given below:

$$f(\lambda) = \begin{cases} \dfrac{1}{\lambda^2} & \text{if } \lambda \geqq 1, \\ 0 & \text{if } 0 < \lambda < 1. \end{cases}$$

Since each $N^{(\lambda)}$ is strictly stationary, so is N. Since no $N^{(\lambda)}$ has any multiple points, nor does N. We have

$$\begin{aligned} \mu = E\{N[0,1]\} &= \int_0^\infty \lambda f(\lambda)\, d\lambda = \int_1^\infty \frac{1}{\lambda^2} d\lambda = +\infty; \\ \frac{r_{k+1}(t)}{t} &= \int_1^\infty \left\{1 - e^{-\lambda t} \sum_{j=0}^{k} \frac{(\lambda t)^j}{j!}\right\} \frac{1}{t\lambda^2} d\lambda. \end{aligned}$$

Making the change of variable $t\lambda = u$, we obtain

(14) $$\frac{r_{k+1}(t)}{t} = \int_t^\infty \left\{1 - e^{-u} \sum_{j=0}^{k} \frac{u^j}{j!}\right\} \frac{du}{u^2}.$$

As $t \downarrow 0$ the limit λ_{k+1} is therefore just the integral in (14) with $t = 0$. Thus it is clear that $0 < \lambda_{k+1} < \infty$ for $k \geqq 1$, but

$$\lambda_1 = \int_0^\infty \frac{1 - e^{-u}}{u^2} du = +\infty. \tag{15}$$

Perhaps the point of the theorems by Korolyuk and Dobrushin is the equality

$$\mu = \lambda_1. \tag{16}$$

There is no reason to expect anything of the sort when multiple points are allowed, as in compound Poisson processes. Thus the two theorems together settle the case where $\mu < \infty$ and there are no multiple points. The case $\mu = \infty$ could be facilely dismissed by applied probabilists as "possessing no practical interest" (see N. B. at the end of this paper). Nevertheless, let us point out that as a corollary to Propositions 4 and 5, (16) is also true when $\mu = \infty$, and for some $k \geqq 2$ there are no points of multiplicity $\geqq k$. For then by (5) and Khintchine's theorem we have

$$\infty = \lambda_1 + \lim_{t \downarrow 0} \sum_{j=2}^{k} \frac{r_j(t)}{t} \tag{17}$$

and the last limit must be finite if $\lambda_1 < \infty$, since r_j decreases as j increases. Thus $\lambda_1 = \infty = \mu$. Another case where this is so is given in the example above. Leadbetter [3] has shown that if there are no multiple points, then $\mu = \infty$ implies $\lambda_1 = \infty$.

3. Extension to several dimensions. We turn now to the consideration of theorems of the above type when R is replaced by R^d, the Euclidean space of dimension d. There are recent studies of point processes in which the points belong to a more general topological space, but so far as I am aware these are not relevant to the questions at hand. The extensions to R^d decidedly possess practical interest, since scientists do count particles with a grid under the microscope, etc. As no new difficulty arises when $d \geqq 3$, we shall take $d = 2$.

Call I an **interval** in R^2 iff it is a bounded parallelogram with its sides parallel to the coordinate axes, but may or may not include all its boundary. Denote its side lengths by $a(I)$ and $b(I)$, its area and diameter by

$$|I| = a(I)b(I), \qquad d(I) = \sqrt{a(I)^2 + b(I)^2},$$

and put

$$\rho(I) = \frac{b(I)}{a(I)}.$$

For each ρ, where $0 < \rho < \infty$, the family of such intervals with $\rho(I) = \rho$ will be denoted by $\mathscr{K}(\rho)$, for example, when $\rho = 1$ these are squares. The family of all intervals will be denoted by $\mathscr{K} = \bigcup_{0<\rho<\infty} \mathscr{K}(\rho)$.

Under assumptions analogous: o (A) and (B), the process $\{N(I,\omega)\}$ with $I \in \mathscr{K}$, $\omega \in \Omega$, will be called a crudely stationary counting process on R^2 iff whenever I and J are two closed intervals of the same area, the random variables $N(I)$ and $N(J)$ have the same distribution. Analogues of Propositions 1 and 2 then hold, but be careful: the intervals I_n in the analogue of (2) must be assumed to be uniformly bounded, in other words, contained in a fixed interval. We have as the analogue to (4), for each interval I:

$$E\{N(I)\} = \mu|I|, \quad \text{where } \mu = E\{N(Q)\}, \tag{18}$$

and Q is a unit square in $\mathscr{K}$.

Fix ρ and a member J of $\mathscr{K}(\rho)$. Then all members of $\mathscr{K}(\rho)$ are congruent to tJ for some $t > 0$, where tJ is an interval homothetic to J at the ratio $t:1$, so that $|tJ| = t^2|J|$. If we now restrict ourselves to members of $\mathscr{K}(\rho)$, we may put

$$\begin{aligned} r_k(t) &= P\{N(tJ) \geqq k\}, \\ \lambda_k(\rho) &= \lim_{t \downarrow 0} \frac{r_k(t)}{t^2|J|}, \end{aligned}$$

whenever the limit exists. Then Khintchine's theorem as well as Propositions 3, 4 and 5 can all be extended to this case. For instance, we have the following trivial extension of the well-known subadditivity lemma used by Khintchine.

LEMMA. *Let ϕ on $[0, \infty)$ be non-negative and have the following property: whenever $0 < t \leqq ns$, where n is a positive integer, we have*

$$\phi(t) \leqq n^2\phi(s).$$

Then we have

$$\lim_{t \downarrow 0} \frac{\phi(t)}{t^2} = \sup_{t > 0} \frac{\phi(t)}{t^2} \leqq \infty.$$

If we set $\phi(t) = r_1(t)$, then Boole's inequality and crude stationarity imply that ϕ satisfies the conditions of the lemma, from which the extended Khintchine theorem follows. Similarly, the proofs of the other propositions carry over to the present case without any difficulty.

However, it is more interesting to consider the larger family $\mathscr{K}$ of all intervals. We then define for $k \geqq 1$:

$$\lambda_k = \lim_{\substack{d(I) \to 0 \\ I \in \mathscr{K}}} \frac{P\{N(I) \geqq k\}}{|I|} \tag{19}$$

whenever the limit exists. The methods used above can be modified to prove the cited propositions in the new context. Everything depends on the following elementary covering lemma:

LEMMA. *Let $I \in \mathscr{K}$ and $\varepsilon > 0$ be given; there exists $\delta = \delta(I, \varepsilon) > 0$ with the following property: For any $J \in \mathscr{K}$ with $d(J) < \delta$, we can find J_j, $1 \leqq j \leqq l$, which are disjoint (apart from sides) and all congruent to J, and which satisfy*

$$I \subset \sum_{j=1}^{l} J_j \subset (1+\varepsilon)I.$$

The proof is omitted as geometrically obvious.

We now state and prove the theorems by Khintchine, Dobrushin and Korolyuk, leaving the previous extensions of the last two theorems to the reader.

PROPOSITION 7. *The limit λ_1 always exists, $\leqq \infty$.*

Proof. Denote by λ_1' the lower limit on the right side of (19), when $k = 1$. Let J_n be a sequence of intervals achieving this lower limit; thus

$$\lambda_1' = \lim_{n \to \infty} \frac{P\{|J_n| \geqq 1\}}{|J_n|}.$$

Since $d(J_n) \to 0$, we may apply the covering lemma to I and J_n for all n such that $d(J_n) < \delta$. Thus we have J_{nj}, $1 \leqq j \leqq l_n$, all congruent to J_n such that

$$I \subset \bigcup_{j=1}^{l_n} J_{nj} \subset (1+\varepsilon)I. \tag{20}$$

It follows from the first inclusion and Boole's inequality that

$$\{N(I) \geqq 1\} \subset \bigcup_{j=1}^{l_n} \{N(J_{nj}) \geqq 1\};$$

and consequently by crude stationarity

$$P\{N(I) \geqq 1\} \leqq l_n P\{N(J_n) \geqq 1\}.$$

On the other hand, the second inclusion in (20) implies that

$$l_n |J_n| \leqq (1+\varepsilon)^2 |I|.$$

Combining the last two inequalities, we obtain

$$\frac{P\{N(I) \geqq 1\}}{|I|} \leqq (1+\varepsilon)^2 \frac{P\{N(J_n) \geqq 1\}}{|J_n|}. \tag{21}$$

'Letting $n \to \infty$ in (21) and then $\varepsilon \to 0$, we see that the left member of (21) does not exceed λ_1'. Since I is arbitrary, this means

$$\sup_{I \in \mathscr{K}} \frac{P\{N(I) \geqq 1\}}{|I|} \leqq \lambda_1';$$

the more so if the "sup" above is replaced by the upper limit as $d(I) \to 0$. Therefore λ_1 exists.

PROPOSITION 8. *If there are no multiple points and $\mu < \infty$, then $\lambda_2 = 0$. If $\lambda_2(\rho) = 0$ for some $\rho > 0$, then almost surely there are no multiple points.*

Proof. For every J in $\mathscr{K}$, we put

$$\xi(J) = l_{[N(J) \geqq 2]}.$$

Let I and J_n be given in $\mathscr{K}$, where $d(J_n) \to 0$; as in the preceding proof, we have (20) for large n. Now define

$$\eta_n(I) = \sum_{j=1}^{l_n} \xi(J_{nj}).$$

If there are no multiple points, then just as in the proof of Proposition 3,

$$P\{\lim_{n \to \infty} \eta_n(I) = 0\} = 1.$$

Since $\eta_n(I) \leqq N((1+\varepsilon)I)$ there is dominated convergence so that

$$\lim_{n \to \infty} E\{\eta_n(I)\} = 0.$$

But from (20) and crude stationarity

$$E\{\eta_n(I)\} \geqq l_n E\{\xi(J_n)\} \geqq |I| \frac{P\{N(J_n) \geqq 2\}}{|J_n|}.$$

Hence

$$\lim_{n \to \infty} \frac{P\{N(J_n) \geqq 2\}}{|J_n|} = 0.$$

This being true for any sequence J_n in $\mathscr{K}$ with $d(J_n) \to 0$, we have $\lambda_2 = 0$.

Conversely, suppose $\lambda_2(\rho) = 0$. Choose any J from $\mathscr{K}(\rho)$ and divide it into 4^n disjoint J_{nj} all congruent to $2^{-n}J$. (This is nothing but Weierstrass' bisection argument.) Clearly

$$\{\xi(J) > 0\} \subset \bigcup_{j=1}^{4^n} \{\xi(J_{nj}) > 0\};$$

hence

$$\begin{aligned} P\{N(J) \geqq 2\} &= P\{\xi(J) > 0\} \leqq 4^n P\{\xi(2^{-n}J) > 0\} \\ &= |J| \frac{P\{N(2^{-n}J) \geqq 2\}}{|2^{-n}J|}. \end{aligned}$$

The last term tends to zero by hypothesis, and J is arbitrary; it follows that there is no multiple point (with probability one).

PROPOSITION 9. *If $\lambda_2(\rho) = 0$ for some $\rho > 0$, then $\lambda_1 = \mu \leqq \infty$.*

Proof. We have remarked that the extension of Korolyuk's theorem is easy for the family $\mathscr{K}(\rho)$. Hence if $\lambda_2(\rho) = 0$ then $\lambda_1(\rho) = \mu$. But by proposition 7, $\lambda_1(\rho) = \lambda_1$ for every ρ.

In conclusion, we may ask what family of figures satisfies a covering property as stated in the Lemma, or some weaker form of it which will still serve the purpose. If we confine ourselves to polygonal ones, then one family is that of all such figures which can be used to *pave* the plane, such as triangles and honeycomb-like hexagons (not necessarily regular) as well as our family $\mathscr{K}$. Paving figures with curved boundaries may be considered provided the boundaries are smooth enough. On the other hand, disks seem to be out, despite Vitali's covering theorem. Nevertheless, are there appropriate extensions of the results discussed here to such figures as disks?

N.B. It is not a mere flight of rhetoric to say that in many mathematical questions, one must ponder over the infinite in order fully to comprehend the finite. Surely the most celebrated instance of this in the history of probability is the St. Petersburg Paradox dealing with the law of large numbers when the mathematical expectation is infinite. A similar situation is the central limit theorem under Lindeberg's condition, when the variance is infinite. Perhaps more relevant to the subject of this note is the existence of quasi-stationary distribution in a recurrent Markov chain, when the steady state must be described by an infinite total mass. This plays a basic role in the deeper parts of the theory. The possibility of infinitely many jumps in finite time, corresponding to the case where $P\{N(t) = +\infty\} > 0$, in the notation of this note is the origin of modern boundary theory, which ought to find applications in various explosive or rapidly changing phenomena. Applied mathematicians are all too apt to dismiss a somewhat delicate situation as pathological or impractical simply because their tools are too crude to cope with them, and then justify this on spurious grounds. It is by no means clear that Nature operates on finiteness assumptions, otherwise why are there infinitely many primes?

Added in proof: I am indebted to Daley and Vere-Jones for the remark that in $R^1, P\{1 \leqq N(t) \leqq k\}$ is subadditive in t, hence λ_k exists for $k \geqq 2$ provided $\lambda_1 < \infty$. A similar result holds for the λ_k defined in (19) by a simple modification of the proof of Proposition 7. See also a forthcoming paper by R. K. Milne in *The Annals of Mathematical Statistics*.

Research supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF under AFOSR Contract F44620-67-C-0049.

References

1. H. Cramér and M. R. Leadbetter, Stationary and Related Stochastic Processes, Wiley, New York, 1967.

2. Y. A. Khintchine, Mathematical Methods in the Theory of Queueing, Griffin, London, 1960.

3. M. R. Leadbetter, On three basic results in the theory of stationary point processes, Proc. Amer. Math. Soc., 19 (1968) 115–117.

THE IMAGE OF THE MATHEMATICIAN

C. V. NEWSOM, Retired Vice-President, RCA

During recent correspondence with Henry Alder, Secretary of the Association, I expressed concern over the apparent fact, as I have been able to observe the educational-industrial-governmental scene, that the image of the professional mathematician as held by American society has undergone serious deterioration in the last few years. If my observation is valid, the future demand for persons with degrees in mathematics will probably be depressed to an even lower level than that which has been anticipated. Professor Alder informed Professor John W. Brace, Chairman of the Committee on the Exchange of Information on Mathematics, of my concern. The following letter, slightly edited and published here at the request of the editor, was written as a reply to a letter which I received from Professor Brace.

Dear Professor Brace:

I appreciated your letter of March 16, for, as Henry Alder has informed you, I have become greatly concerned by the nature of the image of the mathematician presently held by many members of American society. Since writing my original letter to Henry, an article has appeared in the *New Yorker*, written by Alfred Adler, that contains such sentences as the following: "Mathematicians are often expected to manage brilliantly in the fields of business and finance. Of course, they do nothing of the kind. Their non-mathematical efforts are, on the whole, pitifully inept. The qualities embedded in the mind of the mathematician by the discipline of mathematics fail to extend beyond the boundaries of mathematics." Such comments, I must emphasize, represent a very common point of view; thus one hears the question often repeated, "Unless a person expects to teach mathematics, why should he study courses in mathematics beyond the most elementary?" Such a question is given support by the scientist who says, "I learned my mathematics in my courses in science," and by the industrialist who says, "I do not know what to do with a mathematician after I employ him, for it has been my experience that he is unable to isolate and frame the problems that are to be solved." Henry Alder sent me, as you indicate, a copy of the report of the Committee of which you are the Chairman. The thesis of that report, as further stated in your letter of March 16, is that American mathematics faces a serious problem in communication.

Carroll Newsom received his Michigan Doctorate under Walter B. Ford. He held a position at the Univ. of New Mexico, and a Professorship at Oberlin College before becoming head of the Higher Education System in New York State, then Vice-President and President of New York University. He also was the President of Prentice-Hall and Vice-President of R.C.A. He is presently very active in communication science. Dr. Newsom served on the War Policy Committee of Mathematicians, chaired a Committee to Reorganize the MAA, was Editor of the AMERICAN MATHEMATICAL MONTHLY, and was active in curriculum reform. He holds 23 Honorary Doctor's Degrees, has served on over 30 corporate Boards, and is a Board Member of the Guggenheim Foundation. His research interests are function theory, foundations of mathematics, and communication science. He has published numerous books and articles on mathematics, education, and television. *Editor.*

I do not doubt that a well-conceived program of communications would be beneficial to the prestige of American mathematics. But, based on my personal experience, I must advance the hypothesis that poor communication is not a very important element of the problem with which we are presently concerned. It is my judgment that the reputation of mathematics as a fundamental academic discipline is presently being questioned in a way that most mathematicians do not seem to realize; in fact, I believe that it is urgent that causes of the questioning be determined as carefully as possible so that appropriate corrective actions may be taken.

The image of the mathematician that is becoming current, in many of its characteristics and probably in many of its causes, represents a throwback to the situation that existed in the twenties and early thirties. Probably not many people in present-day mathematics are even aware of those days when there was a strong trend in American education to eliminate the mathematics requirement for high school graduation; even in college there was a pronounced deemphasis of mathematics. I attended meetings of engineering educators in those days when great support was given to the idea that all mathematics for engineering students should be taught as a part of the engineering courses. Vigorous steps had to be taken to counter the trend, and the Association was active in working with the National Council of Teachers of Mathematics and with other agencies in trying to understand the reasons for the low status of mathematics. It was decided that new approaches and new emphases were required for the mathematics courses of the schools and colleges. As a result of studies that were undertaken and of the efforts that were initiated, the prestige of mathematics as an academic discipline soon began to make some recovery from the low days of the twenties and the early thirties. Then World War II was upon us, and it became known that two mathematicians, John von Neumann and Stan Ulam, made the two most significant contributions to the work at Los Alamos. So, shortly mathematics had regained status as a basic discipline, perhaps the most basic discipline, for any person who would be truly educated in any science and in many other areas. Then, as we all know, mathematics, like the sciences, profited from the modification in educational priorities stimulated by the appearance of Sputnik.

Now another change has taken place. As indicated above, we seem to be returning to the situation that existed forty years ago; public recognition of mathematics as a fundamental field of study has lost much ground. I suspect that there are two reasons for the reversal in public opinion; at least, I am suggesting two possible reasons as a basis for further consideration. First, John von Neumann and Stan Ulam came out of an academic environment that was different from that of the present-day Ph. D's. It was John who wrote, "The most vitally characteristic fact about mathematics is, in my opinion, its quite peculiar relationship to the natural sciences, or more generally, to any science which interprets experience on a higher than purely descriptive level." John, with whom I had many conversations, could not separate mathematics from life; he saw mathematics wherever he looked. His

feel for nature inspired him to be a better mathematician and his mathematics inspired him to better understand nature. Essentially the same point of view in regard to the relationship between mathematics and nature was inherent in ideas commonly expressed by Richard Courant. But, most present-day recipients of the Ph. D. in mathematics look at their subject as merely a formal discipline without any relevance to nature; moreover, in general they possess no feel for mathematics as a part of our culture or as a factor in the development of our culture. Only a few weeks ago, the head of a large research laboratory expressed his dismay that so few people who had specialized in mathematics had any serious background in a physical science, in economics, in business, or in any other area where mathematics has become important. Then he said, "We are living in an interdisciplinary world. Too many mathematicians have separated themselves from that world."

In the second place, the new elementary mathematics curricula developed in recent years for school and college are superb when analyzed with respect to their mathematical content. They were well designed to produce good mathematicians. I must confess my early satisfaction in regard to the programs. Now, however, we are learning that good mathematicians had too free a hand in the development of the programs. The words of Felix Klein, wise mathematician and pedagogue, were ignored; he wrote: "The presentation (of mathematics) in the schools should be psychological and not systematic. The teacher, so to speak, must be a diplomat. He must take account of the psychic processes in the boy in order to grip his interest; and he will succeed only if he presents things in a form intuitively comprehensible. A more abstract presentation will be possible only in the upper classes." Some students are stimulated by the new programs, but, unfortunately, our educational institutions must deal with a vast number of students, often very competent students, for whom the new programs are only slightly compatible with their interests, their special abilities, and their cultural background. An even more critical situation exists in some of the sciences where the new programs, especially on the secondary level, are killing off the interest of a large number of students; yet, many of those students who are dropping out of the study of science have an innate interest in the subject that would be stimulated of a different approach were employed. It is my judgment that the time has come for a thorough reexamination of the new curricula in the light of actual student interests and needs and the sociology of the day.

The very fact that I have gone into such great detail in this letter reveals the depth of the concern of a person who has lived a varied and complex life but always as a mathematician.

Yours sincerely,

C. V. Newsom

Commentary on the above letter by its author:

As implied in the letter, the experiences of many teachers with the new elementary curricula of school and college seem to demonstrate the validity of the warning of Felix Klein. And another factor that must be recognized in the teaching of science and mathematics has become important in recent years, actually since the philosophy underlying the new curricula was developed; I refer to the changed attitudes and interests of the students. Recently Melvin Kranzberg* has written:

"We must remember that approximately one-third of American college-age youngsters are now in college, and their aversion to required science courses would seem to manifest a disregard or even disrespect for science. Students now are concerned with the quality of life, and they wish to participate more actively in society. Above all, they are motivated by humane and social considerations. Science education has not responded satisfactorily to changing motivations."

Mathematics, with its vast history of relevance to great human accomplishment, can be presented to inquisitive students with their present-day attitudes in a way that will be meaningful to even the most skeptical. Mathematical knowledge has been the underlying factor in providing man many explanations of phenomena with which he has had a concern. And, of very great importance, mathematical thinking as typified by the construction and use of "mathematical models", although such terminology may not be employed and a particular model may provide only a very rough fit to a situation, has become fundamental in a great variety of studies in a diversity of disciplines. Moreover, the successful use of the "systems approach", which many men in various types of endeavor now profess to employ but few understand, is readily accomplished by the person who has had some experience in applying mathematics and has an understanding of the nature of mathematical systems.

The previous letter refers to the efforts of the Association and the National Council of Teachers of Mathematics to resolve some of the critical problems for mathematics that had arisen during the thirties. The efforts involved important contributions by many individuals, especially by some dedicated and enlightened secondary teachers and by a number of professors in small but distinguished colleges; coordination of the efforts was maintained through extensive exchanges of information and materials and by means of continuing discussions by committees. Several of the *Yearbooks* of the National Council provided invaluable assistance to secondary teachers, and even college teachers, in the presentation of new ideas and suggestions, and the sectional meetings of the Association became an important vehicle for the presentation and discussion of ideas of possible significance for the development of instructional programs that would be more meaningful for a majority of college students.

---

* Melvin Kranzberg, "Scientists: The Loyal Opposition", *American Scientist*, January-February, 1072, pp. 20–23.

Popular criticisms of mathematics during the thirties were actually part of a widespread and militant movement for educational reform; it was argued that the schools and colleges had adapted their academic programs in the several disciplines to the needs of potential specialists and had generally ignored the fact that for most students the study of such programs was undertaken for the purpose of providing breadth of understanding. Ultimately, therefore, there was created in the United States, with substantial financial support from a foundation, an elaborate project known as the Cooperative Study in General Education, which sponsored a series of working conferences at the University of Chicago. The project provided an additional affiliation for concerned mathematicians who previously had received the backing of the Association but no financial assistance; the treasury of the Association was under as much pressure then as it is now. The new relationship with the Cooperative Study proved to be fortuitous, for the mathematicians involved in the study, while retaining their previous close contact with the Association and its program, now were involved in well-organized discussions and in planning with college administrators, with college board members, with physical and social scientists, and with assorted consultants. It was decided that in all elementary mathematics courses in college more attention should be given to the foundations and fundamental concepts of mathematics, so that there would be a better understanding of mathematics as a central part of knowledge, without however overwhelming the students with vocabulary and with unduly rigorous mathematical treatments. The admonition of Felix Klein was to be observed. And, of very great importance, as revealed by the perspective of history, it was recommended that, insofar as possible, general formulas should be derived in a mathematics course only after students had experienced the nature of the derivation through the solution of real problems taken from the cultural background of the student. So, in many colleges students were soon indulging in some very sophisticated mathematics as an inductive outgrowth of working with a variety of problems that were meaningful to them. Illustrations of the "new way to mathematics", as it was then called, were presented to many gatherings of educators and to non-educators. The new way provided students experiences with good mathematics, but, in addition, it continually made them aware of the fact that mathematics is an intimate part of man's life. Unfortunately, the extensive program of developing appropriate instructional materials that was to take place after there had been a sufficient amount of classroom experience with the new ideas was virtually terminated by the advent of World War II.

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

ON AN INEQUALITY OF J. W. S. CASSELS

RALPH ALEXANDER, University of Illinois

If $z_1, z_2, \cdots, z_n$ are complex numbers in the disc $\{z : |z| \leqq 1\}$, we have the classical inequality

$$\prod_{i \neq j} |z_i \bar{z}_j - 1| \leqq n^n, \qquad \text{(see [2])}. \tag{1}$$

Equality holds in (1) only if the z_i are the vertices of a regular n-gon inscribed in the unit circle. The following theorem of Cassels [1] generalizes (1).

THEOREM 1. (Cassels) *Let $z_1, \cdots, z_n$ be complex numbers in the disc $\{z : |z| \leqq \rho, \rho > 1\}$. Suppose $\cos(\pi/n) \leqq \rho^2/(\rho^4 - \rho^2 + 1)$. Then we have the inequality*

$$\prod_{i \neq j} |z_i \bar{z}_j - 1| \leqq (\rho^{2n} - 1)^n (\rho^2 - 1)^{-n}. \tag{2}$$

Equality holds only if the z_i are vertices of a regular n-gon inscribed in the circle of radius ρ.

The following corollary suffices for the applications made by Professor Cassels.

COROLLARY 1. (Cassels) *Let $z_1, \cdots, z_n$ be complex numbers in the disc $\{z : |z| \leqq \rho, \rho > 1\}$. Suppose $\rho \leqq 1 + (1/10n)$. Then we have the inequality*

$$\prod_{i \neq j} |z_i \bar{z}_j - 1| \leqq n^n \rho^{2n(n-1)}. \tag{3}$$

We conjecture that (2) remains true without the condition on $\cos(\pi/n)$. While we were not able to establish this, we can give a corresponding improvement on Corollary 1.

If $P(z, \omega)$ is a polynomial in two complex variables, let $M(P, n)$ denote $\max \prod_{i \neq j} |P(z_i, z_j)|$, as $z_1, \cdots, z_n$ ranges over all the numbers on the circle, $|z| = 1$.

LEMMA 1. *Let $P(z, \omega)$ be a homogeneous polynomial of degree s. Let $z_1, \cdots, z_n$ and $\omega_1, \cdots, \omega_n$ be numbers in the disc, $|z| \leqq \rho$. Then*

$$\prod_{i \neq j} |P(z_i \bar{z}_j, \omega_i \bar{\omega}_j)| \leqq \rho^{2n(n-1)s} M(P, n). \tag{4}$$

Proof. Let z_1 be considered as a variable while holding $z_2, \cdots, z_n$ and $\omega_1, \omega_2, \cdots, \omega_n$ fixed. We note that for each $j \neq 1, \left| P(z_j\bar{z}_1, \omega_j\bar{\omega}_1) \right| = \left| \bar{P}(z_1\bar{z}_j, \omega_1\bar{\omega}_j) \right|$. Hence the norm of the product on the left side of (4) agrees with the norm of a polynomial in z_1. Applying the maximum modulus principle, the product will achieve a maximum for some $z_1, \left| z_1 \right| = \rho$. The same argument is repeated for each z_i and ω_i. Let us assume that the norm of each number equals ρ. Since P is homogeneous of degree s,

$$P(z_i\bar{z}_j, \omega_i\bar{\omega}_j) = (\bar{z}_j\omega_i)^s P(z_i/\omega_i, \bar{\omega}_j/\bar{z}_j).$$

However, $\bar{\omega}_j/\bar{z}_j = z_j/\omega_j$ since $\left| z_j/\omega_j \right| = 1$. Note that the left side of (4) will consist of $n(n-1)$ factors whose product will be at most $\rho^{2n(n-1)s} M(P, n)$.

THEOREM 2. *Let z_i, w_i $(i \leqq n)$ lie in the disc, $\left| z \right| \leqq \rho$. Then*

$$\prod_{i \neq j} \left| z_i\bar{z}_j - \omega_i\bar{\omega}_j \right| \leqq \rho^{2n(n-1)} n^n. \tag{5}$$

Equality holds only if the numbers z_i/ω_i are the vertices of a regular n-gon inscribed in the circle $\left| z \right| = \rho$.

Proof. We let $P(z, \omega) = z - \omega$. It follows immediately from (1) that $M(P, n) = n^n$. The inequality follows upon application of Lemma 1.

If $\rho \geqq 1$, we may choose $\omega_i = 1$ for each i, and observe that the conclusion of Corollary 1 follows without assuming that $\rho \leqq 1 + (1/10n)$.

As a final remark, we claim that a careful analysis of Professor Cassels' proof of Theorem 1 allows us to make a slight improvement to the effect that $\cos(\pi/n) \leqq 2\rho^2/(\rho^4 + 1)$ suffices.

I wish to thank Professor Kenneth B. Stolarsky for stimulating discussions which led to this note.

References

1. J. W. S. Cassels, On a problem of Schinzel and Zassenhaus, J. Math. Sciences, 1(1966) 1–8.

2. I. Schur, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, Math. Z., 1(1918) 377–402.

SETS WHICH SPLIT FAMILIES OF MEASURABLE SETS

R. B. KIRK, Southern Illinois University

The purpose of this note is to formulate and to prove for abstract measure spaces a generalization of the feature of Lebesgue measure contained in the following problem of W. Rudin ([3], page 56). "Construct a Borel set $E \subset R^1$ such that $0 < m(E \cap I) < m(I)$ for every non-empty segment I." The proof of the generalization affords a nice application of the Baire category theorem. In order to state the result,

we make the following definition. Let $(X, \mathscr{S}, m)$ be a measure space and let $\mathscr{G}$ be a subset of $\mathscr{S}$ consisting of sets of positive measure. A set $A \in \mathscr{S}$ **splits** $\mathscr{G}$ if $0 < m(A \cap B) < m(B)$ for all $B \in \mathscr{G}$. Recall that the measure m is **atomless** if $A \in \mathscr{S}$ and $0 < m(A) \leqq \infty$ imply that there is a $B \in \mathscr{S}$ such that $B \subset A$ and $0 < m(B) < m(A)$. The main result may be stated as follows:

THEOREM 1. *Let $(X, \mathscr{S}, m)$ be a measure space where m is atomless. If $\mathscr{G}$ is a countable subset of $\mathscr{S}$ consisting of sets of positive measure, then there is a set $A \in \mathscr{S}$ which splits $\mathscr{G}$.*

Before presenting the proof we shall give an application to show that it is indeed a generalization of the property of Lebesgue measure noted above.

COROLLARY. *Let X be a separable metric space and let m be an atomless Borel measure on X. Then there is a Borel set A such that $0 < m(A \cap B) < m(B)$ for every open set B of positive measure.*

Proof. Let $\{x_n : n \in N\}$ be dense in X and let Q^+ denote the set of positive rationals. By Theorem 1 there is a Borel set A of finite measure which splits $\mathscr{G} = \{B^0(x_n; r) : n \in N,\ r \in Q^+ \text{ and } 0 < m(B^0(x_n; r)) < \infty\}$. (Of course, $B^0(x_n; r)$ denotes the open ball about x_n of radius r.) It is clear that A also splits the family of open sets of positive measure and the proof is complete.

We shall now proceed with the task of proving the theorem. Let $(X, \mathscr{S}, m)$ be a measure space. When m is finite, define $d(A,B) = \int |\mathscr{K}_A - \mathscr{K}_B|\, dm$ where $\mathscr{K}_A$ and $\mathscr{K}_B$ denote the characteristic functions of A and B. If A and B are identified when $d(A,B) = 0$, then $(\mathscr{S}, d)$ is a complete metric space. (See [**2**], pages 168 and 169.) A set $A \in S$ has the *Darboux property* if $0 < m(A) < \infty$ and if for every real number a with $0 \leqq a < m(A)$, there is $B \in \mathscr{S}$ with $B \subset A$ and $m(B) = a$.

LEMMA. *Let $(X, \mathscr{S}, m)$ be a measure space with m finite. For $B \in \mathscr{S}$ define $F(B) = \{A \in \mathscr{S} : m(A \cap B) = 0 \text{ or } m(A \cap B) = m(B)\}$. Then $F(B)$ is a closed subspace of $(\mathscr{S}, d)$. Furthermore, if B has the Darboux property, then $F(B)$ is nowhere dense in $(\mathscr{S}, d)$.*

Proof. For $A_1, A_2 \in \mathscr{S}$, note that

$$\mathscr{K}_{A_1 \cap B} - \mathscr{K}_{A_2 \cap B} \leqq |\mathscr{K}_{A_1} - \mathscr{K}_{A_2}|$$

and hence that $|m(A_1 \cap B) - m(A_2 \cap B)| \leqq d(A_1, A_2)$. Thus the function $A \to m(A \cap B)$ is continuous from $(\mathscr{S}, d)$ to R. Since $F(B)$ is the inverse image of $\{0, m(B)\}$ for this function, $F(B)$ is closed.

Now assume that B has the Darboux property. Take $A \in F(B)$ and $\varepsilon > 0$ arbitrarily. We shall show that $B^0(A; \varepsilon) - F(B) \neq \varnothing$. Indeed, since B has the Darboux property, there is a set $E \in \mathscr{S}$ such that $E \subset B$ and $0 < m(E) < \min(\varepsilon, m(B))$. Since $A \in F(B)$, there are two cases to consider.

Case 1. Assume that $m(A \cap B) = 0$. Define $A^* = A \cup E$. Note that

$$0 < m(A^* \cap B) \leqq m(E) < m(B),$$

even though $d(A, A^*) = m(A^* - A) = m(E - A) \leqq m(E) < \varepsilon$.

Case 2. Assume that $m(A \cap B) = m(B)$. Define $A^* = A - E$. Note that

$$0 < m(A^* \cap B) = m(B - E) = m(B) - m(E) < m(B),$$

even though $d(A, A^*) = m(A - A^*) = m(A \cap E) < m(E) < \varepsilon$.

Thus whichever case is applicable, the set A^* defined satisfies $A^* \in B^0(A, \varepsilon) - F(B)$.

We now state a theorem which will be needed in the proof of Theorem 1. A proof of this theorem may be found in [1], page 26.

THEOREM 2. *Let $(X, \mathscr{S}, m)$ be a measure space where m is atomless. If $B \in \mathscr{S}$ is such that $0 < m(B) < \infty$, then B has the Darboux property relative to m.*

Proof of Theorem 1. First assume that m is finite and let $\mathscr{G} = \{B_1, B_2, \cdots\}$. For each $n \in N$, B_n has the Darboux property by Theorem 2. Hence by the lemma, $F(B_n)$ is a closed nowhere dense subset of $(\mathscr{S}, d)$ for each $n \in N$. Since $(\mathscr{S}, d)$ is complete, we conclude from Baire's category theorem that there is a set $A \in \mathscr{S} - \cup\{F(B_n): n \in N\}$. It is clear that A splits $\mathscr{G}$.

In general, for each $n \in N$, choose $E_n \subset B_n$ such that $0 < m(E_n) < \min(2^{-n}, m(B_n))$. (This is possible by Theorem 2.) Define $E = \cup\{E_n: n \in N\}$. Let m_E denote the restriction of m to E. (That is, $m_E(A) = m(A \cap E)$ for all $A \in \mathscr{S}$.) Then m_E is finite and so there is $A \in \mathscr{S}$, such that $0 < m_E(A \cap E_n) < m_E(E_n)$ for all n. That is, $0 < m(A \cap E_n) < m(E_n)$ for all n. It is then clear that A splits $\mathscr{G}$. The proof is complete.

References

1. N. Dinculeanu, Vector Measures, Pergamon, New York, 1967.
2. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, N. J., 1950.
3. W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1966.

REPRESENTATIVES FOR COSETS

JAMES ALONSO, Bennett College, Greensboro, N. C.

It is well known that if H is a subgroup of finite index, or finite order of a group G, then there exists a common set of representatives for the left and the right cosets of H in G. For finite index see [8] pages 12 and 37. For finite order ([1] and [2]) it is sufficient to see that each double coset $C = HxH$ contains the same number of left and of right cosets of H and that each left coset of H in C meets each right coset of H in C, in fact the left coset h_1xH and the right coset Hxh_2 have h_1xh_2 in common.

We cannot say the same for tee case of a subgroup H of infinite index and infinite order; two counterexamples can be found in [**6**] and [**10**] Ex. 9.2.12 page 218.

More generally, if H and K are subgroups of the same finite index ([**9**] Th. 4.3. and [**11**]) or of the same finite order (see [**5**]) of a group G, then there exists a common set of representatives for the left cosets of H and the right cosets of K. The proposition cannot be generalized for subgroups of the same infinite index and the same infinite order; a trivial counterexample is given by the additive group of rational numbers and the subgroups of integers and of even integers. Ore [**9**] gives other conditions under which common sets of representatives exist for the left cosets of a subgroup and the right cosets of another.

PROPOSITION 1. *If H and K are subgroups of the same finite index or the same finite order of a group G, then there exists a common set of representatives for the right cosets of H and the right cosets of K.*

In the case of finite index, the proposition can be proven by applying to a set of representatives of the right cosets of the subgroup $H \cap K$ the following theorem due to König [**3**]:

If a set is divided in a finite number m of disjoint classes in two different ways and r classes of the first subdivision contain at most r classes of the second, then the two subdivisions have a common set of representatives.

In the case of finite order, the proposition follows immediately from the following theorem of König-Valkó [**4**] and van der Waerden [**6**], which also applies to the case of right cosets of H and left cosets of K:

If a set is divided in two different ways in disjoint classes of the same finite number n of elements, then the two subdivisions have a common set of representatives.

The theorems of König and König-Valkó–van der Waerden are particular cases of a more general proposition due to De Bruijn [7].

In the particular case of subgroups H and K of order 2, the proof of proposition 1 contained virtually in [7] can be significantly simplified in the following way:

Define the following relation in G: xRy if there exists a finite chain $A_1, A_2, \cdots, A_n$ of right cosets alternately of H and K such that $x \in A_1$, $y \in A_n$ and A_i meets A_{i+1} for $i = 1, \cdots, n-1$. R is obviously an equivalence relation, and each equivalence class is disjoint union of right cosets of H and disjoint union of right cosets of K. In each equivalence class C choose arbitrarily one element x_0; call the rest $x_1, x_2, \cdots, x_{-1}, x_{-2}, \cdots$, with x_{2i} in the same right coset of H with x_{2i+1} and in the same right coset of K with x_{2i-1} for each integer i. Take for representatives the elements $x_0, x_2, \cdots, x_{-2}, x_{-4}, \cdots$. The class C may be an infinite chain (fig. 1) or a closed finite chain of cosets (fig. 2); in either case each right coset of H or K n C obtains a unique representative.

(In figs. 1 and 2 the points represent elements of C, the straight segments cosets of H, the curved segments cosets of K, the arrows are used to point out the elements taken as representatives for cosets.)

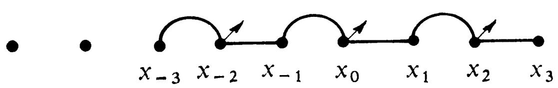


FIG. 1

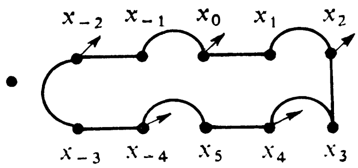


FIG. 2

PROPOSITION 2. *If G is the free product of the two non-trivial subgroups H and K, then there exists a common set of representatives for the right cosets of H and the right cosets of K in G.*

Proof. Each right coset of H distinct from H can be expressed in a unique way in the form

$$\bar{H} = Hx_1x_2\cdots x_n \tag{1}$$

for a positive integer n with $x_i \neq 1$ for $i = 1,2,\cdots,n$ and $x_i \in H$ if i is even, $x_i \in K$ if i is odd. Conversely, each expression of the form (1) yields a right coset of H. A similar statement holds for the right cosets of K. We say that the coset $\bar{H} = Hx_1x_2\cdots x_n$ has length n. Write

$$C_n = \{\bar{H} \mid \bar{H} \text{ is a right coset of } H \text{ of length } n\},$$

$$D_n = \{\bar{K} \mid \bar{K} \text{ is a right coset of } K \text{ of length } n\},$$

$n = 1,2,\cdots$. $C_0 = \{H\}$, $D_0 = \{K\}$.

(i) if $\bar{H} \in C_n$ $(n > 0)$ then $\bar{H}$ meets exactly one coset of K in D_{n-1}, the rest of the cosets of K met by $\bar{H}$ are in D_{n+1} and their cardinal number is $o(H) - 1$; in fact $Hx_1x_2\cdots x_n$ meets the cosets $Kx_2\cdots x_n$ and $Khx_1x_2\cdots x_n$ $(1 \neq h \in H)$.

(ii) if $\bar{K}$ and $\bar{\bar{K}}$ are right cosets of K in D_{n+1} that meet distinct cosets of H in C_n then $\bar{K}$ and $\bar{\bar{K}}$ are distinct; in fact $\bar{K} = Ky_1y_2\cdots y_{n+1}$ and $\bar{\bar{K}} = Kz_1z_2\cdots z_{n+1}$ meet $\bar{H} = Hy_2\cdots y_{n+1}$ and $\bar{\bar{H}} = Hz_2\cdots z_{n+1}$ respectively; hence if $\bar{H} \neq \bar{\bar{H}}$ then $\bar{K} \neq \bar{\bar{K}}$.

Statements similar to (i) and (ii) can be made changing the roles of H and K. Therefore we can find a common set of representatives in the following way:

(1) Take 1 for representative of H and K.

(2) Take arbitrarily representatives for the cosets of H in C_1 and for the cosets of K in D_1. These will also represent distinct cosets of K in D_2 and distinct cosets of H in C_2.

(3) Take arbitrarily representatives for those cosets of H in C_2 and for those cosets of K in D_2 which still have no representative.

(4) Repeating the preceding process, every coset of H and every coset of K obtains a unique representative.

Example 1. Let G be the group of 2×2 regular matrices over the field R of rational numbers; H the subgroup of diagonal matrices in G. The existence of a common system of representatives for the left and the right cosets of H in G is insured by Ore ([9] Th. 2.1). The reader can easily verify that the following is one such system:

$$\left\{\begin{pmatrix}1 & r\\ 0 & 1\end{pmatrix} \,\middle|\, r \in R\right\} \cup \left\{\begin{pmatrix}r & 1\\ 1 & 0\end{pmatrix} \,\middle|\, r \in R\right\} \cup \left\{\begin{pmatrix}0 & 1\\ 1 & r\end{pmatrix} \,\middle|\, 0 \neq r \in R\right\} \cup \left\{\begin{pmatrix}1 & 0\\ r & 1\end{pmatrix} \,\middle|\, 0 \neq r \in R\right\}$$

$$\cup \left\{\begin{pmatrix}r & 1\\ 1 & t\end{pmatrix} \,\middle|\, 0 \neq rt \neq 1;\ r, t \in R\right\}.$$

Example 2. If $F[x, y]$ is the free group in two generators, it can be seen that the words of the form

$$x_1 y_1 x_2 y_2 \cdots x_n y_n x_{n+1},$$

where the x's are powers of x and the y's powers of y, none of them the zero power, except perhaps x_1, together with the empty word form a common set of representatives for the left and the right cosets of the subgroup $\langle x \rangle$ in $F[x, y]$. (The existence of such a common set of representatives follows from [9] Th. 2.1.)

Example 3. Let G be the group of rigid motions of the plane under composition, H and K the subgroups of order 3 generated respectively by the rotations of 120° around two different points A and B. A common set of representatives can be found for the right cosets of H and the right cosets of K in the following manner:

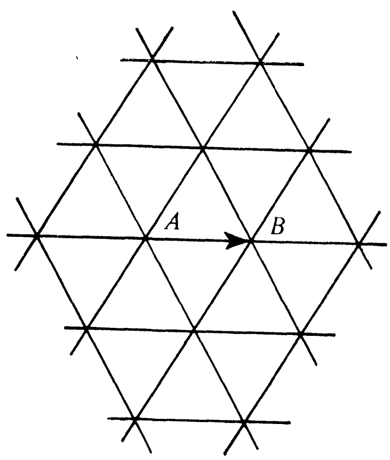


FIG. 3

Call L the subgroup of G of the rigid motions that take the arrow $\overrightarrow{AB}$ to one of the edges of the infinite lattice based on AB (fig. 3). Each right coset of H contained in L has one element and only one that takes the arrow $\overrightarrow{AB}$ to a parallel

position, and the same holds for the right cosets of K in L. Hence the set

$$T = \{t \mid t \in L \text{ and } t \text{ takes } \overrightarrow{AB} \text{ to a parallel position}\}$$

is a common set of representatives for the right cosets of H and the right cosets of K in L. Now if X is a set of representatives for the right cosets of L in G, $\{tx \mid t \in T \text{ and } x \in X\}$ is the desired set.

References

1. G. A. Miller, On a method due to Galois, Quart, J. Math. Oxford Ser. 41 (1910) 382–384.

2. H. W. Chapman, A note on the elementary theory of groups of finite order, Messenger of Math., 42 (1913) 132–134.

3. D. König, Über Graphen und ihre Anwendungen auf Determinantentheorie und Mengenlehre, Math. Ann., 77 (1916) 453–465.

4. D. König and St. Valkó, Über mehrdeutige Abbildungen von Mengen, Math. Ann., 95 (1926) 135–138.

5. G. Scorza, A proposito di un teorema di Chapman, Boll. Un. Mat. Ital., 6 (1927) 1–6.

6. B. L. van der Waerden, Ein Satz über Klasseneinteilungen von endlichen Mengen, Abh. Hamb. Sem., 5 (1927) 185–188.

7. N. G. De Bruijn, Gemeenschappelijke representantensystemen van twee klassenindeelingen van een verzameling, Nieuw Arch. Wisk. (ser. 2), 22 (1943) 48–52.

8. H. Zassenhaus, The Theory of Groups, Chelsea, New York, 1949.

9. O. Ore, On coset representatives in groups, Proc. Amer. Math. Soc., 9 (1958) 665–670.

10. W. R. Scott, Group Theory, Prentice Hall, Englewood Cliffs, N. J., 1964.

11. E. Weiss, Coset representatives, Portugaliae Mathematica, 26 (1967) 259–260.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

HOW TO CUT ALL EDGES OF A POLYTOPE?

BRANKO GRÜNBAUM, University of Washington

In the sequel we shall discuss a family of mutually related problems which are somewhat remarkable from two points of view: On the one hand, despite their intuitive appeal and accessibility, the questions seem to have been considered in the literature only in very few isolated instances. On the other hand, the ramifications of the topic reach from pattern recognition (which motivated the investigation of

O'Neil [7]) through the theories of graphs and of convex polytopes to functional analysis (which provided the motivation for Klee [6]).

Let P be a d-polytope (that is, a convex polytope of dimension d in Euclidean d-space E^d; for terminology and results concerning polytopes see [2]); a **cut** of P is any set of edges of P which may be simultaneously intersected by a $(d-1)$-dimensional hyperplane that misses all the vertices of P. We define the **cut-number** $m(P)$ of P as the minimal number of cuts needed to cover all the edges of P. For a trivial example we may take $d = 2$ and P any n-gon; then clearly $m(P) = [\frac{1}{2}(n+1)]$, where $[x]$ is the largest integer not exceeding x.

If $k \leqq d$ then k hyperplanes divide E^d into at most 2^k regions. Therefore, denoting by T^d any d-dimensional simplex, it is obvious that $m(T^d) =]\log_2(d+1)[$ where $]x[$ denotes the smallest integer not less than x. This may be generalized to the following conjecture which, although trivial for $d = 2, 3$, is open for all $d \geqq 4$.

CONJECTURE 1. *The cut-number of every d-polytope P satisfies*

$$m(P) \geqq m(T^d) =]\log_2(d+1)[.$$

If C^d denotes the regular d-dimensional cube, it is easily checked that $m(C^3) = 3$ and $m(C^d) \leqq d$. Recently O'Neil [7] proved that any cut of C^d contains at most

$$[\tfrac{1}{2}(d+1)] \binom{d}{[\frac{1}{2}d]}$$

edges; therefore

$$m(C^d) \geqq d2^{d-1}/[\tfrac{1}{2}(d+1)] \binom{d}{[\frac{1}{2}d]}$$

which, using Stirling's formula, leads to a bound of the type $m(C^d) \geqq an^{\frac{1}{2}}$ for a suitable constant $a > 0$. It may be verified that $m(C^4) = 4$, but the reasonably seeming conjecture $m(C^d) = d$ is reported by O'Neil to be false, an example attributed to Paterson implying $m(C^6) \leqq 5$.

This leads to

Problem 1. Determine $m(C^d)$ for $d \geqq 5$.

Denoting by Q^d the (regular) d-dimensional cross-polytope, it is easy to verify that $m(Q^d) \leqq 1 +]\log_2 d[$, and it may be conjectured that equality holds for all d. Moreover, we make

CONJECTURE 2. *The cut-number of every centrally symmetric d-polytope P satisfies* $m(P) \geqq 1 +]\log_2 d[$.

At least for $d = 3$ Conjecture 2 is true and we have (see Grünbaum [3]):

For every centrally symmetric 3-*polytope* P *we have* $m(P) \geqq 3$.

As an analogue of Conjecture 1 for simple polytopes we make:

CONJECTURE 3. *If P is a simple d-polytope then* $m(P) \geqq m(C^d)$.

We also venture

CONJECTURE 4. *If P and P' are isomorphic d-polytopes then*

$$m(P) = m(P').$$

This conjecture appears almost preposterous, in view of the fact that the maximal number of edges in a cut may differ for isomorphic polytopes. For example, any cut of the regular octahedron has at most 6 edges, while the xy-plane determines an 8-edge cut in the isomorphic polytope with vertices $(\pm 1, 0, \frac{1}{4})$, $(0, \pm 1, -\frac{1}{4})$, $(0,0,\pm 1)$. The chances of the conjecture being true are naturally better if only simple d-polytopes are considered, or if $d = 3$—but even if both conditions are satisfied the proof seems to be very elusive. In this context we make:

CONJECTURE 5. *The maximal number of edges in cuts of every polytope isomorphic to the d-cube C^d is the same as for cuts of C^d itself.*

If true this conjecture is of interest since polytopes isomorphic to C^d allow cuts of many types not possible with C^d. For example, the convex hull of the 8 points $(\pm 3, \pm 1, 1)$ and $(\pm 1, \pm 3, -1)$ in E^3 is isomorphic to the cube C^3; the plane $x = y$ intersects it in six edges that correspond to the heavily drawn edges of C^3 in Figure 1—but the only 6-edge cuts of C^3 form a circuit of length 6.

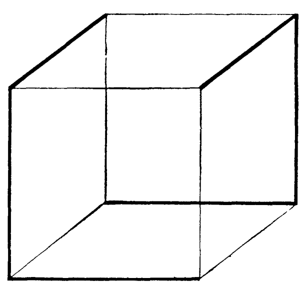


FIG. 1.

Let us call an ***i*-cut** of P ("i" for "isomorphism") a set of edges of P which correspond to the edges of a cut in some polytope P' isomorphic to P, and let $m_i(P)$ be the least number of i-cuts needed to cover all edges of P. Clearly $m_i(P) \leqq m(P)$, and strict inequality holds even for C^3. Indeed $m_i(C^3) = 2$, since the heavy lines in Figure 1 form one i-cut of C^3, while the thin ones form another. Clearly, $m_i(T^d) = m(T^d)$ for all d; we have

Problem 2. Determine $m_i(C^d)$ for $d \geqq 4$.

CONJECTURE 6. *For every d-polytope P,* $m_i(P) \geqq m_i(T^d)$; *moreover, if P has a center of symmetry then* $m_i(P) \geqq m_i(C^d)$.

Instead of d-polytopes it is possible to consider **tessellations** of the $(d-1)$-sphere S^{d-1}, that is, cell-complex decompositions of the unit sphere S^{d-1}, with convex cells. While each d-polytope leads by radial projection to such tessellations, it is well known (see Supnick [**9**], Shephard [**8**]) that not every tessellation is obtainable as such a projection. If the cut-number $m(Q)$ of a tessellation Q of S^{d-1} is defined as the least number of great $(d-2)$-spheres (which miss all vertices of Q) needed to intersect all edges of Q, one may reformulate for tessellations Conjecture 1. However, the analogue of Conjecture 2 has to be modified since (see Grünbaum [**3**]) there exist centrally symmetric tessellations Q of the 2-sphere with $m(Q) = 2 < m(C^3) = 3$. This leads to:

Problem 3. Determine $\min_Q\{m(Q)\}$, where Q ranges over all centrally symmetric tessellations of the $(d-1)$-sphere.

We call a t-**cut** ("t" for "topological") of a d-polytope P any set of edges intersectable by a suitable homeomorphic image S of a $(d-2)$-sphere S^{d-2} in boundary of P, provided $F \cap S$ is a topological j-cell, $j \leqq k-1$, for every k-face F of P. We define $m_t(P)$ as the least number of t-cuts needed to cover all edges of P. Clearly $m_t(P) \leqq m_i(P)$, and we make

CONJECTURE 7. *For every* $d \geqq 4$ *there exists a d-polytope* P_d *such that* $m_t(P_d) < m_i(P_d)$.

In some contrast to that conjecture is the following fact:

For every 3-*polytope P we have* $m_i(P) = m_t(P)$.

In order to prove this assertion it is clearly enough to show that every t-cut of a 3-polytope is also an i-cut. Let P be a 3-polytope, let H be a t-cut of P, and let P^* be a polytope dual to P. Then there is a natural one-to-one correspondence between the edges of P and those of P^*, such that to edges of each t-cut H of P there correspond the edges of a simple circuit H^* in P^*, and vice versa. According to a beautiful theorem of Barnette [**1**], there exists a 3-polytope $\bar{P}$ isomorphic to P^* such that the circuit $\bar{H}$ which corresponds to H^* is the (sharp) **shadow boundary** of $\bar{P}$ for projection (illumination) in the direction of a suitable line L. Denoting by P' a polytope polar to $\bar{P}$ with respect to a sphere centered at an interior point of $\bar{P}$, and by H' the t-cut of P' that corresponds to the circuit $\bar{H}$ of $\bar{P}$, it follows that the plane through the origin perpendicular to L intersects all the edges in H'; hence H' is a cut, and our assertion is proved.

Because of the duality between the t-cuts of 3-polytopes and simple circuits on the dual polytopes, it is possible to deduce some properties of $m_t(P)$ from known results on simple circuits (see [**4**] for a survey of results and for references). As an example we mention the following fact:

There exist 3-polytopes P with arbitrarily many edges, having all vertices of valence $\leqq 6$ and all faces with at most 6 sides, such that $m_t(P) \geqq (e(P))^b$, where b is a positive constant (we may even take $b \geqq 1-(\log 8/\log 13) \approx 0.19$), and $e(P)$ is the number of edges of P.

We could not decide:

Problem 4. Does there exist a constant b such that $m(P) \leqq b$ (or at least $m_t(P) \leqq b$) for every 3-polytope P having only vertices of valence at most 4 and faces with at most 4 sides?

The ideas discussed above may be modified to apply to graphs. Let G be a graph; a ***g*-cut** of G (cocircuit; cocyle in Harary [**5**]) is a set of edges of G which separates G and is minimal with respect to that property (i.e., no proper subset separates). We define $m_g(G)$ to be the least number of g-cuts needed to cover all edges of G. Clearly, if G is the graph of a d-polytope P, then every t-cut of P is a g-cut of G. In case $d=3$ the converse is also true, but for $d \geqq 4$ it is not known whether every g-cut of the graph of a d-polytope is a t-cut of the polytope. Analogues of Conjectures 1 and 2 may be formulated for g-cuts of the graphs of d-polytopes.

The properties of graphs of d-polytopes lead naturally to some problems concerning g-cuts of graphs.

Problem 5. Determine $m_g(k)$, the least value of $m_g(G)$ when G varies over all k-connected graphs.

It is not hard to verify, using the graphs of the tetrahedron, octahedron, and icosahedron, that $m_g(3) = m_g(4) = m_g(5) = 2$, but the value of $m_g(6)$ is not known. However, we make

CONJECTURE 8. *If G is a k-connected graph and if G contains a subgraph isomorphic to a subdivision of the complete graph with $k+1$ nodes, then $m_g(G) \geqq \log_2(k+1)$.*

The various problems posed above may also be generalized in a completely different direction. Let P be a d-polytope, and let k be an integer with $1 \leqq k \leqq d-2$. We call a ***(k)-cut*** of P any set of k-faces of P which may be simultaneously intersected by a suitable $(d-k)$-flat H such that $H \cap F = \emptyset$ for all faces F of P of dimension less than k. The definition of (k)-cut-number of P, etc., is obvious. The cuts we discussed above correspond to $k=1$. Unfortunately, no non-trivial results on those notions seem to be known for $k>1$, except that a result of Klee [**6**] may be interpreted as follows:

Every centrally symmetric d-polytope has a $(d-2)$-cut comprising at least $2d$ $(d-2)$-faces.

Added in proof: The validity of Conjecture 1 has been established by David W. Barnette; his paper "Cut numbers of convex polytopes" will appear in the journal *Geometriae Dedicata.*

Research supported in part by the Office of Naval Research under Grant N00014-67-A-0003.

References

1. D. W. Barnette, Projections of 3-polytopes, Israel J. Math., 8 (1970) 304–308.
2. B. Grünbaum, Convex Polytopes, Interscience, New York, 1967.
3. ———, Intersecting all edges of centrally symmetric polyhedra by planes, (to appear).
4. B. Grünbaum and H. Walther, Shortness exponents of families of graphs (to appear).
5. F. Harary, Graph Theory, Addison-Wesley, Reading, Mass., 1969.
6. V. Klee, On a conjecture of Lindenstrauss, Israel J. Math., 1 (1963) 1–4.
7. P. E. O'Neil, Hyperplane cuts of an n-cube, Discrete Math., 1 (1971) 193–195.
8. G. C. Shephard, Spherical complexes and radial projections of polytopes, Israel J. Math., 9 (1971) 257–262.
9. F. Supnick, On the perspective deformation of polyhedra, Ann. Math., 49 (1948) 714–730, and 53 (1951) 55–555.

CORRECTIONS TO "THE HADAMARD MAXIMUM DETERMINANT PROBLEM"

(This MONTHLY, 79(1972) 626–630.)

JOEL BRENNER AND LARRY CUMMINGS

Please note the following:

1. The correct address of J. L. Brenner is: 10 Phillips Rd. Palo Alto, CA 94303.
2. The research was supported by NSF GP 32527.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

A UNIFIED PROOF OF SEVERAL BASIC THEOREMS OF REAL ANALYSIS

PATRICK SHANAHAN, College of the Holy Cross, Worcester, Massachusetts.

1. Introduction. The place of continuity in elementary real analysis is justified by its role as a hypothesis in three important theorems. Specifically, if f is a continuous real-valued function on a closed bounded interval $[a, b]$, then

(i) f is bounded on $[a, b]$ (and actually attains maximum and minimum values);
(ii) f has the intermediate value property on $[a, b]$; and
(iii) f is Riemann integrable on $[a, b]$.

It is the purpose of this note to present proofs of these theorems in which the part played by continuity is isolated and shown to enter into each proof in essentially the same way; in effect, the three theorems are derived as corollaries of a single lemma.

2. The main lemma. Let $[a, b]$ be a given closed, bounded interval in the real number system, with $a \leqq b$. Let $\mathscr{C}$ be a family of subsets of $[a, b]$. Let us say that $\mathscr{C}$ is **local** if each point $x \in [a, b]$ has a neighborhood, with respect to the relative topology on $[a, b]$, which is a member of $\mathscr{C}$. Say that $\mathscr{C}$ is **additive** if whenever C_1 and C_2 are members of $\mathscr{C}$ such that $C_1 \cap C_2 \neq \varnothing$, then $C_1 \cup C_2 \in \mathscr{C}$.

LEMMA 1. *If $\mathscr{C}$ is a local, additive family of closed subintervals of $[a, b]$, then $[a, b] \in \mathscr{C}$.*

Proof. Let $D = \{x \mid [a, x] \in \mathscr{C}\}$. We wish to prove that $b \in D$. Since $\mathscr{C}$ is local, there is an interval $[a, y]$ in $\mathscr{C}$, and therefore $D \neq \varnothing$. Since each member of $\mathscr{C}$ is contained in $[a, b]$, D is bounded. Thus D has a least upper bound d.

Since d is the least upper bound of D, every neighborhood of d meets D. In particular, taking a neighborhood $[d', d'']$ of d which belongs to $\mathscr{C}$, we see that there is an element $d_0 \in D$ such that $d' \leqq d_0 \leqq d$. Then $[a, d_0] \in \mathscr{C}$, and thus by the additivity of $\mathscr{C}$ we have

$$[a, d''] = [a, d_0] \cup [d', d''] \in \mathscr{C},$$

which means that $d'' \in D$, and hence that $d = d''$.

In other words, $[d', d]$ is a neighborhood of d relative to the interval $[a, b]$. But this can happen only if $d = b$. Thus $b = d = d'' \in D$, which completes the proof.

3. Proof of Theorems 1, 2, and 3.

THEOREM 1. *If f is continuous on $[a, b]$, then f is bounded on $[a, b]$.*

Proof. Let $\mathscr{C}$ consist of all closed subintervals of $[a, b]$ on which f is bounded.

$\mathscr{C}$ is local. For, given $x \in [a, b]$ the continuity of f at x implies that there exists a neighborhood $[c, d]$ of x relative to $[a, b]$ such that for all $y \in [c, d]$ we have $|f(y) - f(x)| < 1$. Thus f is bounded on $[c, d]$ and hence $[c, d] \in \mathscr{C}$.

$\mathscr{C}$ is additive. For, if f is bounded on the closed intervals C_1 and C_2, it is bounded on $C_1 \cup C_2$. If $C_1 \cap C_2 \neq \varnothing$ then $C_1 \cup C_2$ is again a closed interval, and hence $C_1 \cup C_2 \in \mathscr{C}$.

Applying Lemma 1, we have $[a, b] \in \mathscr{C}$. That is, f is bounded in $[a, b]$. (It follows easily from this that f actually attains a maximum and a minimum value on $[a, b]$.)

THEOREM 2. *If f is continuous on $[a, b]$, then f has the intermediate value property on $[a, b]$.*

Proof. It is sufficient to show that if f takes on both positive and negative values on $[a, b]$, then it must have a zero in $[a, b]$.

Suppose f has no zeroes in $[a, b]$. Let $\mathscr{C}$ consist of all closed subintervals of $[a, b]$ on which the sign of f is constant.

$\mathscr{C}$ is local. For, given $x \in [a, b]$, the continuity of f at x implies that there exists a neighborhood $[c, d]$ of x relative to $[a, b]$ such that for all $y \in [c, d]$ we have

$|f(y) - f(x)| < |f(x)|$. The inequality implies that $f(y)$ and $f(x)$ have the same sign, that is, the sign of f is constant on $[c, d]$. Thus $[c, d] \in \mathscr{C}$.

$\mathscr{C}$ is additive. For, let f have constant sign on the closed intervals C_1 and C_2. If $x \in C_1 \cap C_2$, then the sign of f on each interval agrees with the sign of $f(x)$, and hence f has constant sign on the closed interval $C_1 \cup C_2$. Therefore $C_1 \cup C_2 \in \mathscr{C}$.

Applying Lemma 1, we have $[a, b] \in \mathscr{C}$, that is, f has constant sign on $[a, b]$. But this contradicts the assumption that f takes on both positive and negative values on $[a, b]$.

THEOREM 3. *If f is continuous on $[a, b]$, then f is Riemann integrable on $[a, b]$.*

Proof. By Theorem 1, f is bounded on $[a, b]$, and therefore the upper and lower integrals $\overline{\int_c^d} f$ and $\underline{\int_c^d} f$ exist for any closed interval $[c, d] \subset [a, b]$. Let $\varepsilon > 0$ be given. It is enough to show that

$$\overline{\int_a^b} f - \underline{\int_a^b} f \leqq (b - a)\varepsilon,$$

from which it follows that $\overline{\int_a^b} f = \underline{\int_a^b} f$. (We assume the elementary properties of upper and lower Riemann integrals.)

Let $\mathscr{C}$ consist of all closed subintervals $[c, d]$ of $[a, b]$ which have the property that for any interval $[c', d'] \subset [c, d]$

$$\overline{\int_{c'}^{d'}} f - \underline{\int_{c'}^{d'}} f \leqq (d' - c')\varepsilon.$$

$\mathscr{C}$ is local. For, given $x \in [a, b]$, the continuity of f implies that there is a neighborhood $[c, d]$ of x such that the difference $M - m$ between the maximum value M and the minimum value m of f on $[c, d]$ is less than ε. For any subinterval $[c', d'] \subset [c, d]$ we then have

$$\overline{\int_{c'}^{d'}} f - \underline{\int_{c'}^{d'}} f \leqq (d' - c')M - (d' - c')m \leqq (d' - c')\varepsilon.$$

Hence $[c, d] \in \mathscr{C}$.

$\mathscr{C}$ is additive. For, let $[c_1, d_1]$ and $[c_2, d_2]$ be members of $\mathscr{C}$ with non-empty intersection. We may assume that neither is contained in the other, and that in fact $c_1 < c_2 \leqq d_1 < d_2$. Now let $[c', d']$ be a subinterval of $[c_1, d_1] \cup [c_2, d_2]$. Either $[c', d']$ is contained in one of the terms of the union, in which case there is nothing to prove, or we must have $c' < c_2 < d'$. In the latter event, since $[c', c_2] \subset [c_1, d_1]$ and $[c_2, d'] \subset [c_2, d_2]$, it follows that

$$\begin{aligned}\overline{\int_{c'}^{d'}} f - \underline{\int_{c'}^{d'}} f &= \overline{\int_{c'}^{c_2}} f + \overline{\int_{c_2}^{d'}} f - \underline{\int_{c'}^{c_2}} f - \underline{\int_{c_2}^{d'}} f \\ &\leqq (c_2 - c')\varepsilon + (d' - c_2)\varepsilon = (d' - c')\varepsilon.\end{aligned}$$

Thus $[c_1, d_1] \cup [c_2, d_2] \in \mathscr{C}$.

Applying Lemma 1, we have $[a, b] \in \mathscr{C}$, which implies in particular that

$$\overline{\int_a^b} f - \underline{\int_a^b} f \leqq (b-a)\varepsilon.$$

REMARK: As an alternative method of proof, one could use Lemma 1 to prove that f is uniformly continuous on $[a, b]$ (take $\mathscr{C}$ to be the family of closed subintervals on which f is uniformly continuous) and then apply a standard partition refinement argument.

4. Generalizations. It follows easily from Lemma 1 that closed intervals are compact. More generally, if in the definition of a local additive family of subsets given in Section 2 we replace the closed interval $[a, b]$ by an arbitrary topological space X, the following proposition holds:

PROPOSITION 1. *Let X be a non-empty connected topological space. Then the following statements are equivalent:*

(i) *X is compact,*

(ii) *X is a member of every local, additive family of subsets of X.*

Proof. Assume that X is compact, and let $\mathscr{C}$ be a local, additive family of subsets of X. Since $\mathscr{C}$ is local, the interiors $\mathring{C}$ of the members of $\mathscr{C}$ constitute an open covering of X. Since X is compact, a finite subcollection $\mathring{C}_1, \mathring{C}_2, \cdots, \mathring{C}_n$ of these sets covers X. Since X is non-empty and connected, we may assume that these sets are ordered so that $(C_1 \cup C_2 \cup \cdots \cup C_{k-1}) \cap C_k \neq \varnothing$ for $1 < k \leqq n$. By the additivity of $\mathscr{C}$ we then have $X = C_1 \cup C_2 \cup \cdots \cup C_n \in \mathscr{C}$.

Conversely, assume that X satisfies condition (ii) and let $\mathscr{U}$ be an open covering of X. Define $\mathscr{C}$ to be the family of all subsets of X which are contained in the union of a finite number of sets of $\mathscr{U}$. Since every $x \in X$ belongs to some member of $\mathscr{U}$, the family $\mathscr{C}$ is local. If C_1 and C_2 can each be covered by a finite number of sets of $\mathscr{U}$, then so can $C_1 \cup C_2$, hence $\mathscr{C}$ is additive. Thus $X \in \mathscr{C}$, that is, X is compact.

A slight generalization in another direction is possible. Let us call a family $\mathscr{C}$ of subsets of X **sub-additive** if whenever C_1 and C_2 are members of $\mathscr{C}$ such that $C_1 \cap C_2 \neq \varnothing$, then $C_1 \cup C_2$ is contained in some member of $\mathscr{C}$.

PROPOSITION 2. *Let X be a topological space. Then the following statements are equivalent:*

(i) *X is a member of every local, sub-additive family on X,*

(ii) *X is a member of every local, additive family on X.*

Proof. Since additive families are sub-additive, (i) implies (ii). Assume that X satisfies (ii), and let $\mathscr{C}$ be a local, sub-additive family on X. Define $\mathscr{C}'$ to be the family of all subsets of X which are contained in some member of $\mathscr{C}$. It is clear that $\mathscr{C}'$ is local and additive. By (ii), $X \in \mathscr{C}'$. But this means that X is contained in a member of $\mathscr{C}$, that is, $X \in \mathscr{C}$.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

THE CHINESE MATHEMATICAL OLYMPIADS: A CASE STUDY

FRANK SWETZ, Capitol Campus Pennsylvania State University

In 1960, the MONTHLY contained an article by John De Francis describing the inauguration of a Mathematical Olympiad in the Peoples Republic of China [**1**]. The article provided a fascinating glimpse of the Chinese State's efforts to nurture mathematical talent among its youth. Since that time, little information has reached the West concerning either Chinese mathematics education or the fate of the Olympiad scheme. Chinese Mathematical Olympiads continued up until 1964 and achieved many of the goals they were designed for: mathematically talented students were located and given special educational attention; the general level of mathematics instruction was elevated and thousands of Chinese students were encouraged to come together in "study groups" for extra curricular mathematics studies. In light of the recent decision to institute a Mathematical Olympiad in the United States, it might prove beneficial to examine the Chinese experience, its execution and its consequences.

The Conception and Execution of the Chinese Mathematical Olympiads. During a visit to the Soviet Union in April of 1946, Hua Lo-keng, Director of the Institute of Mathematics of the Chinese Academy of Science, was impressed by the enthusiastic response given by secondary school students to a lecture on complex numbers by P. S. Aleksandrov of Moscow University. These students were members of study groups preparing themselves for participation in the Soviet Mathematical Olympiads. Returning to Russia in 1953, with a delegation from *Academica Sinica*, Hua and his colleagues were advised by Soviet educators to instigate mathematical competitions in China as a method of promoting scientific advancement. The consensus was that through such activities Chinese youth would be stimulated toward mathematics studies thus forcing an improvement in the quality of school mathematics and science instruction. Firmly convinced of the potential national benefits of mathematical competitions, Hua suggested their adoption in the Peoples Republic of China as an extra-curricular activity but also cautioned against a resulting disruption in the regular school system [**2**]. The examinations were not to interfere with the school's normal functions. He was supported in this campaign by Tuan Hsueh-fu, professor at Peking University, who urged Chinese educators to "learn from Russia" concerning mathematical competitions [**3**].

Early in 1956, activities began in earnest to implement Hua and Tuan's recommendations. Mathematics competition committees were established in Peking, Shanghai, Tientsin and Wuhan. They were responsible for local organization of contests and for setting examination questions. Shanghai's committee was composed of seventeen members selected from the Mathematical Society, the Shanghai Municipality Education Office and the local chapter of the Chinese National Association of Natural and Special Sciences. In the choosing of examination questions, the committees limited their selection to topics from arithmetic, algebra, geometry and trigonometry that while rigorous did not exceed content required by middle school mathematics outlines. Similar to the Soviet scheme, associated student lectures on various aspects of mathematics were to be given. The initial lecture in this series was presented on March 11 by Professor Su Pu-chin. His topic was "Non-Euclidean Geometries."

In May the first examination was undertaken. Students in the last two years of middle school were given a screening examination by their teachers. Those who did well and were politically acceptable were recommended to represent their schools in city-wide competition. The official examination was composed of two rounds with the final winners emerging from the second round. Each round contained five or six problems in a given time allotment of one hundred fifty minutes. Students who passed the first round were awarded a certificate of merit and allowed to compete in the second round. Success at the last level warranted a medal and an award of books. The competitors with the three best scores were permitted entrance to the universities of their choice to study either mathematics, physics, astronomy, or any other associated scientific discipline without being subjected to further examinations. Naturally, the accomplishment of doing well in such an examination brought great recognition to the young scholar and for a short period he became a local hero much like the successful civil service candidates of old. On May 4th, Wuhan conducted its examinations and had twenty-one students pass. Sixty-two Peking middle schools sponsored six hundred twenty-two students in the final round of its competition of May 13th. Thirty-three passed. Tientsin's examination on May 27th had four hundred ninety-nine participants in the final round with twenty-five passes. Shanghai's Olympiad was given in early June and saw seven hundred thirty-two contestants in the second round. (No information is available as to the number of final winners [4].) Although the examination efforts in these four cities were considered experimental, they were acclaimed outstanding successes. Shanghai's experiences of 1956 and the following year, 1957, were well documented and published to serve as a guide for other cities to follow [5].

One hundred thirty thousand copies of *Compilation of Problems from the 1956–57 Mathematical Competitions for Middle-School Students in Shanghai Municipality* were published and distributed in 1958. In this booklet the ultimate objectives of the competitions were specified: to locate mathematically talented students so they

could be singled out for special educational attention and to encourage self-study and a competitive spirit among students. Both objectives were intended to raise the quality of mathematical training for Chinese students so that the Peoples Republic of China could compete, scientifically, with the more developed nations of the world.

As a result of the competitions, mathematics study groups were formed in many schools. Students engaged in extra-curricular activities designed to improve their performance on up-coming examinations. Study groups existed on several levels: within schools, among several schools and at the city-wide level. By 1962, the Peking Mathematics Study Group boasted a membership of seven hundred. Members came together once a month to hear a lecture by a prominent mathematician and to engage in discussions concerning his presentation [6]. Often the lecturer would pose specific problems to be solved by his audience. In 1960, the Office of Mathematics, Physics and Logic of the Institute of Mathematics of the Chinese Academy of Science, organized a series of twenty lectures to be presented in future months and designed for student study groups. These lectures centered on four themes:

(1) An introduction to the study of mathematical foundations.

(2) Outline of the history of mathematics.

(3) The nature, methods and significance of mathematics.

(4) The techniques and characteristics of modern mathematics [7].

Eventually, many of the lectures were published in pamphlet form for further and more widespread study by student groups [8].

These lectures and publications were part of a broad government sponsored campaign to promote the study of science. At the forefront of this campaign was Hua Lo-keng. Mathematician of world renown, concerned teacher and confirmed advocate of the Communist Party's policies, Hua was to be emulated as the socialist model of a scholar-scientist. The story of his proletarian background and "Horatio Alger" rise to success despite adversity was communicated to the youth of China with the hope that it would encourage them to be persistent in achieving their educational ideals. The People's Publishers in Shanghai printed his biography, *The Mathematician Hua Lo-keng* and Hua, himself, wrote *To a Young Mathematician* in which he included autobiographical sketches and encouragement to students [9]. Hua was indeed a self-made man worthy of admiration. Although lacking higher academic degrees, he has written several classical works of mathematics, is a versatile researcher and world recognized authority in number theory, harmonic analysis of functions of several complex variables, and group theory [10].

In subsequent years since 1956, the level of achievement on the competitions has increased. This record is due largely to the influence of student mathematics study groups. The 1962 competitions in Peking attracted one thousand four hundred and sixty-five students from one hundred schools, six hundred ninety three seniors,

and seven hundred seventy-two juniors representing five percent of their respective grades city-wide. On the first round nearly half of the seniors scored about 60% correct. The second round was quite difficult but one student did solve all the required problems [11]. Of the eighty-two eventual winners, half were members of the Peking Mathematics Study Group [12]. From data available on the 1963 competitions, it appears that all student participants took both examinations rather than being screened out by the first round.

Peking Municipality Mathematical Competitions

April 12, 1963 (8:00–9:00 A. M. and 9:30–11:30 A. M.) [13]

Junior Level Examination: First Round

1. 10 people are grouped into two clubs, each club consisting of 5 members. In each club a president and a vice-president are chosen. How many ways can this be done?
2. Given: $\sin \alpha + \sin \beta = p$, $\cos \alpha + \cos \beta = q$, find the values of $\sin(\alpha + \beta)$ and $\cos(\alpha + \beta)$.
3. Solve the simultaneous equations:

$$\sqrt{x-1} + \sqrt{y-3} = \sqrt{x+y}$$
$$\lg(x - 10) + \lg(y - 6) = 1.$$

4. The lengths of the sides of a right triangle form three consecutive terms of an arithmetic progression. Prove that the lengths are in the ratio 3:4:5.
5. Let D be a point on the arc $\overset{\frown}{BC}$ of the circumscribed circle about the equilateral triangle ABC. Let E be the intersection of the lines AB and CD, F the intersection of the lines AC and BD. Prove $\overline{BC}$ is the geometric mean of $\overline{BE}$ and $\overline{CF}$. $[\overline{BC}^2 = \overline{BE} \cdot \overline{CF}]$.

Junior Level Examination: Second Round

1. Let $x^3 + bx^2 + cx + d$ be a polynomial with integral coefficients, and let $bd + cd$ be odd. Prove the polynomial is not the product of two polynomials, each with integral coefficients.
2. Suppose 5 points are given in the plane, no 3 on a line, no 4 on a circle. Prove there exists a circle through 3 of the points such that of the remaining 2 points, one is in the interior and the other is in the exterior of the circle.
3. Let P be a point in the interior of a regular hexagon whose sides have length 1. The line segments from P to two vertices have length 13/12 and 5/12 respectively. Determine the lengths of the segments from P to the 4 remaining vertices.
4. Let a be a positive integer, and let $r = \sqrt{a+1} + \sqrt{a}$. Prove that for any positive integer n there exists a positive integer a_n satisfying: $r^{2n} + r^{-2n} = 4a_n + 2$, $r^n = \sqrt{a_n + 1} + \sqrt{a_n}$.

Senior Level Examination: First Round

1. If $2 \lg(x - 2y) = \lg x + \lg y$, find $x{:}y$.
2. Let r and R be the radii respectively of the inscribed and the circumscribed circles to a regular n-gon whose sides have length a. Prove: $r + R = (a/2) \cot \pi/2n$.
3. Find the coefficient of x^2 in

$$(1 + x)^3 + (1 + x)^4 + (1 + x)^5 + \ldots + (1 + x)^{n+2}.$$

4. Given a convex n-gon, call the line segment joining two non-adjacent vertices a diagonal. Assume no 3 diagonals intersect in a common point. Find the number of intersections of diagonals (in the interior of the n-gon).

5. A trapezoid is given with parallel edges of lengths a and $2a$. A side of the trapezoid has length b and forms an acute angle α with the edge of length $2a$. Find the volume of the solid of revolution determined by rotating the trapezoid about the side of length b.

Senior Level Examination: Second Round

1. Let $P(x) = A_kX^k + A_{k-1}X^{k-1} + \cdots + A_1X + A_0$ be a polynomial with integral coefficients. Suppose x_1, x_2, x_3, x_4 are distinct integers such that $P(x_i) = 2$ for $i = 1, 2, 3, 4$. Prove that $P(x)$ is not 1, 3, 5, 7, or 9 for any integer x.

2. Let 9 points be given in the interior of the unit square. Prove there exists a triangle of area $\leqq 1/8$ whose vertices are 3 of the 9 points.

3. $2n + 3$ points are given in the plane, no 3 on a line, no 4 on a circle. Is it possible to find a circle through 3 of the points such that of the remaining $2n$ points, half are in the interior and half are in the exterior of the circle? Prove your answer.

4. 2^n counters are divided into several piles. The following defines a move: choose two piles A and B, say with p and q counters respectively, $p \geqq q$; move q counters from A and put them in pile B. Prove there exists a finite number of moves such that all counters end up in one pile.

The examination was later criticized as being very difficult [14]. Examinations similar to this one were taking place in Peking up until 1964.

The Conclusion and Consequences of the Examination Scheme. It was originally hoped that the mathematical competition schemes would eventually be adopted by all large cities in China. Although the movement did spread from the four cities that inaugurated the tests, it did not achieve the momentum expected. Perhaps in many locales, the mathematical talent and organizational ability for such an endeavor were lacking. The era of "antichampionism" in the sixties and The Great Cultural Revolution terminated the examinations. Under pressure from the red guards Hua had to publicly confess his sin of promoting "advanced experience from abroad" in the Peoples Republic [15]. The competitions were denounced as contributing to elitist education practices by encouraging personal achievement. In the years between 1956 and 1964, the existence of the competitions did much to mould the mathematical thinking patterns of Chinese students. The questions stressed creative thinking over rigid solution methods dictated by rote-learning experiences. Thousands of students benefited from this exposure. Now in the wake of the Great Cultural Revolution, it remains to be seen if the educators in the Peoples Republic of China will consider this fact important enough to resurrect the mathematical competitions.

The author wishes to express his thanks for the kind assistance rendered by Professor Arthur Pu of the University of Illinois at Chicago Circle, in translating the Peking questions.

References

1. John De Francis, Mathematical Competitions in China, this MONTHLY, 67 (1960) 756–762.

2. Hua Lo-keng, We will have National Mathematics Competitions Soon, *Shuxue Tongbao*, Chinese Math. Assoc. Peking, January 1956, pp. 1–3.

3. Tuan Hsueh-fu, Learn from Russia to have Mathematical Competitions, *Shuxue Tongbao*, January 1956, pp. 3–5.

4. Hua Lo-keng, Completion of the Peking Competition, *Shuxue Tongbao*, June 1956, pp. 1–2.

5. Shang-hai shih, 1956–57 nien chung hsüeh-sheng shu-hsüeh ching-sai his-t'i pien-hui, (*Compilation of Problems* from 1956–57 Mathematics Competitions for Middle-School Students in Shanghai Municipality), New Knowledge Press, Shanghai, 1958.

6. Han Erh-Tsai, They Like Mathematics, *China Reconstructs*, Peking, December 1962, 11:34–35.

7. Office of Mathematics, Physics and Logic of the Institute of Mathematics of the Chinese Academy of Science Sponsors Lectures on Mathematics Foundations, *Shuxue Tongbao*, February 1960, p. 42.

8. *Shuxue Tongbao*, September 1962, Back cover. This *Series of Mathematics for Youth* included the following works:

Hua Lo-keng, *Discussions Starting From π of Tsu Ch'ung Discussions Starting from the Triangle of Yang Hui.*

Wu Wen-chun, *Applications of Mechanics in Geometry*

Shih Chi-huai, *Averages.*

Tuan Hsueh-fu, *Symmetry, Induction and Deduction.*

Min Szu-hao, *Lattice Points and Area.*

Chiang K'en-ch'eng, *One Stroke Diagrams and the Mailman's Route.*

Tseng K'en-cheng, *One Hundred Mathematical Problems.*

Ch'ang Keng'che and Wu Jun-sheng, *Complex Numbers and Geometry.*

9. *Shu Hsueh Chia Hua Lo-Keng*, (The Mathematician Hua Lo-keng) People's Publishers, Shanghai, 1956; Hua Lo-keng, *To a Young Mathematician*, China Youth Press, Shanghai, 1956.

10. Some of his publications include: *Additive Prime Number Theory*, Chinese Academy of Sciences, Peking, 1953; *Harmonic Analysis of Functions of Several Complex Variables in Classical Domains*, Izdat. Inostra. Lit., Moscow, 1959; *Classical Groups*, Shanghai Science and Technology Press, 1963 (with Wang Yuan).

11. Conclusions of the 1962 Mathematics Contest Among Middle School Students in Peking Municipality, *Shuxue Tongbao*, April 1963, pp. 50–51.

12. Han, Op. cit., p. 35.

13. 1963 Peking Municipality Mathematical Competitions, *Shuxue Tongbao*, May 1963, back cover; Other published examination questions can be found in: 1957 Peking Municipality Mathematical Competitions, *Shuxue Tongbao*, May 1957, pp. 38–44; 1957 Tientsin, Wuhan and Nanking Mathematical Competitions, *Shuxue Tongbao*, August 1957, pp. 45–46; and *Compilation of Problems from* 1956–57 *Mathematical Competitions for Middle Students in Shanghai Municipality*, Op. cit.

14. Chao Ts'u-keng, On the Problems Adopted for the 1963 Mathematics Contest for Peking Middle School Students, *Shuxue Tongbao*, July 1963, pp. 8–14.

15. Hua Lo-keng, Chairman Mao Points Out the Road of Advance for Me, *China Reconstructs*, November 1969, pp. 30–31 and 41.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before January 31, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2373. *Proposed by Grahame Bennett, Indiana University*

Let $r_1, r_2, \cdots, r_n$ be real numbers. Show that there exists a subset N of $\{1, 2, \cdots, n\}$, neither containing nor omitting three consecutive integers, such that

$$\left| \sum_{j \in N} r_j \right| \geqq \frac{1}{6} \sum_{j=1}^{n} |r_j| .$$

Show further that 1/6 is the best possible constant here.

Establish the corresponding result (with 1/6 replaced by $1/3\pi$) for complex numbers.

E 2374. *Proposed by Judith Q. Longyear, Pennsylvania State University*

Suppose that $a_1 \leqq a_2 \leqq \cdots \leqq a_n$ are natural numbers such that $a_1 + \cdots + a_n = 2n$ and such that $a_n \neq n + 1$. Show that if n is even, then for some subset K of $\{1, 2, \cdots, n\}$ it is true that $\sum_{i \in K} a_i = n$. Show that this is true also if n is odd when we make the additional assumption that $a_n \neq 2$.

E 2375. *Proposed by H. Kestelman, University College, London, England*

Let G be an abelian group. For any subset S of G, let $D(S)$ denote the set of

differences $x - y$, where $x, y \in S$. Show that if A and B are any subsets of G such that $G = A \cup B$, then either $D(A) \supseteq B$ or $D(B) \supseteq A$. Show further that if $G = A \cup B$ and if A and B are not disjoint, then $D(A) = G$ or $D(B) = G$.

E 2376. *Proposed by Arthur Marshall, Madison, Wisconsin*

Suppose that p and q are odd primes and that a and b are natural numbers such that $p^a > q^b$. Show that if p^a divides the product $\sigma(p^a)\sigma(q^b)$, then in fact $p^a = \sigma(q^b)$.

E 2377. *Proposed by Lawrence Somer, University of Illinois*

Find the number of essentially different ways that an element of the finite field $\mathrm{GF}(p^n)$ can be represented as the sum of two squares.

E 2378. *Proposed by D. E. Penney, University of Georgia*

Let a, m and n be natural numbers. Evaluate

$$(a^m + 1,\ a^n + 1).$$

Compare Problem E 2295 [1972, 398].

SOLUTIONS OF ELEMENTARY PROBLEMS

Summations with Ordered Indices

E 2313 [1971, 904]. *Proposed by Sidney Heller, Brookhaven National Laboratory*

Show that

$$\sum_{i_m=1}^{n-m+1} \sum_{i_{m-1}=i_m+1}^{n-m+2} \cdots \sum_{i_3=i_4+1}^{n-2} \sum_{i_2=i_3+1}^{n-1} \sum_{i_1=i_2+1}^{n} 1 = \binom{n}{m}.$$

I. *Solution by Michael Shimshoni, Weizmann Institute of Science, Rehovot, Israel.* We see that $n \geqq i_1 > i_2 > \cdots > i_m \geqq 1$. Any combination of i's satisfying this inequality will appear in the sum once and only once, so that in the sum we have as many summands as there are m-element subsets of $\{1, 2, \cdots, n\}$, viz. $\binom{n}{m}$.

II. *Solution by F. G. Schmitt, Jr., Berkeley, California.* Denoting the given sum by $S(n, m)$ and reversing the order of summation, we have

$$S(n,m) = \sum_{i_1=m}^{n} \sum_{i_2=m-1}^{i_1-1} \sum_{i_3=m-2}^{i_2-1} \cdots \sum_{i_{m-1}=2}^{i_{m-2}-1} \sum_{i_m=1}^{i_{m-1}-1} 1.$$

Obviously $S(n, 1) = n = \binom{n}{1}$; assume as the induction hypothesis that $S(n, m-1) = \binom{n}{m-1}$; then

$$S(n,m) = \sum_{i_1=m}^{n} \binom{i_1 - 1}{m - 1} = \binom{n}{m}.$$

This last summation identity is well known; see, e.g., W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I (Second edition), 1957, p. 61.

III. *Solution by G. P. Steck, Sandia Laboratories, Albuquerque.* Let $k = n - m$ and consider any arrangement of m A's and k B's. Let the position in the sequence of the rth A from the right be i_r so that the rightmost A is in place i_1 and the leftmost A is in position i_m.

The required sum is the number of ways the A's and the B's can be arranged so that $1 \leqq i_m \leqq k+1$, $i_m + 1 \leqq i_{m-1} \leqq k+2, \cdots, i_2 + 1 \leqq i_1 \leqq m + k$. But these restrictions are automatically satisfied since $i_r > i_{r+1}$ and since the rth A from the right cannot be fewer than r places from the right hand end of the sequence. Consequently the required sum is the number of ways that m A's and k B's can be arranged in a sequence, which is $\binom{m+k}{m} = \binom{n}{m}$.

More general sums of the same type appear in ballot problems and in the two-sample problem of order statistics. In this latter context, I have showed that for given sequences of integers $a_1 \leqq a_2 \leqq \cdots \leqq a_m$ and $b_1 \leqq b_2 \leqq \cdots \leqq b_m$, $(i \leqq a_i \leqq b_i \leqq k + i)$, the number of ways m A's and k B's can be arranged so that $a_r \leqq i_{m-r+1} \leqq b_r$ $(r = 1, 2, \cdots, m)$ is the determinant of the $m \times m$ matrix $M = (m_{ij})$ where

$$m_{ij} = \binom{b_i - a_j + j - i + 1}{j - i + 1}.$$

In the case at hand we have $a_i = i$ and $b_i = k + i$. See G. P. Steck, *The Smirnov two sample tests as rank tests*, Ann. Math. Stat. 40 (1969), 1449–1466; a simpler proof is given in S. G. Mohanty, *A short proof of Steck's result on two-sample Smirnov statistics*, Ann. Math. Stat. 42 (1971), 413–414.

Also solved by the proposer and $\binom{10}{2} + 1$ other contributors.

Editor's comment. John Ivie points out that the result can be obtained using generating functions and Pascal's triangle as in his article, *Multiple Fibonacci sums*, Fibonacci Quart. 7 (1969), 303–309. For a connection with lattice problems see C. A. Church and H. W. Gould, *Lattice point solution of the generalized problem of Terquem and an extension of Fibonacci numbers*, Fibonacci Quart. 5 (1967), 59–68. For a connection with Catalan numbers see Problem E 2054 [1969, 192].

Venn Again

E 2314 [1971, 904]. *Proposed by A. K. Austin, The University, Sheffield, England*

Prove or disprove that it is possible to find a convex polygon and three translations of it in the plane which form a Venn diagram for four sets (i.e., they form 16 connected regions and no three edges pass through the same point).

Solution by Heiko Harborth, Braunschweig, Germany. Any two congruent convex polygons that are related by a translation have at most two points of intersection, common arcs being considered as single points. If three such polygons meet in one point, then slight translations of one or two of them will form a small triangle in place of the point, increasing the number of regions by one. Thus we need only consider cases where the polygons intersect two by two in distinct points. A further permissible simplification now is the replacement of the convex polygons by circles. Then the Venn diagram for n such circles has $n(n-1)$ vertices and $2n(n-1)$ edges or arcs ($2n-2$ of each on each circle). By Euler's formula, the number of faces is given by

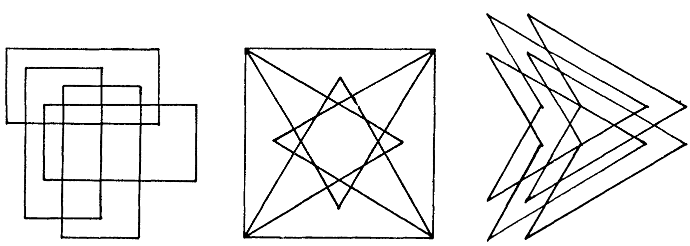
$$F = 2 + E - V = n^2 - n + 2 < 2^n$$

when $n \geqq 4$. Hence a Venn diagram for $n \geqq 4$ sets cannot be formed from any convex set and $n-1$ translations of it.

Also solved by Ken Brons, D. Ž Djoković, J. R. Kuttler, L. E. Mattics, E. T. Ordman, and F. G. Schmitt, Jr.

Editor's comment. Schmitt notes that the proof for circles appears in Yaglom & Yaglom, *Challenging Mathematical Problems with Elementary Solutions*, Vol. I, 1964, 103–104.

Five correspondents sent figures showing four congruent convex polygons (or ovals) forming a Venn diagram and related by translations and rotations. The figures below show such a diagram for rectangles (by G. A. Heuer, Concordia College) and for equilateral triangles (by the reviewer), each of which can be constructed using rotations only. The last figure (by the reviewer) shows four non-convex quadrilaterals related solely by translations. (The last two figures are not connected.)



Subdivisions of a Polygon

E 2315 [1971, 904]. *Proposed by Richard Stanley, Harvard University*

Let $f(n)$ be the number of ways an $(n+1)$-sided convex polygon can be divided into regions by diagonals not intersecting in the interior of the polygon. The trivial division, that is the division using no diagonals, is to be counted, so that $f(1) = 1$, $f(2) = 1, f(3) = 3, f(4) = 11$, etc. Find the generating function $F(x) = \sum f(n)x^n$, and find an asymptotic formula for $f(n)$.

I. *Solution by F. G. Schmitt, Jr., Berkeley, California.* Let P be a convex $(n+1)$-gon with vertices $x_0, x_1, \cdots, x_n$, and let us use the term *diagonal set* to denote a set of diagonals of a polygon which do not intersect in the polygon's interior. (In particular, the null set is a diagonal set.) Then, by definition, $f(n) = |D_n|$, where D_n is the family of all diagonal sets of P. But we can write D_n as the following disjoint union

$$D_n = A_n \cup \bigcup_{k=2}^{n-1} B_{kn},$$

where A_n is the family of diagonal sets of P which do not contain any diagonals through x_0, and where B_{kn} is the family of diagonal sets of P which contain the diagonal x_0x_k but none of the diagonals x_0x_j with $j < k$. Hence, if $a_n = |A_n|$ and $b_{kn} = |B_{kn}|$, we have

$$f(n) = a_n + 2\sum_{k=2}^{n-1} b_{kn}.$$

For $n \geqq 3$, every diagonal set of the convex n-gon $x_1 \cdots x_n$ is in A_n, as is every such set augmented by the inclusion of x_1x_n; moreover, every diagonal set in A_n is of one (and only one) of these two types. The diagonal sets in B_{kn} can be characterized as follows: they contain x_0x_k and are partitioned by it into two independently chosen subsets—one a diagonal set of the convex $(n-k+2)$-gon $x_0x_k \cdots x_n$, and the other a diagonal set of the convex $(k+1)$-gon $x_0 \cdots x_k$ which does not contain any diagonals through x_0. Thus, for $n \geqq 3$, we see that $a_n = 2f(n-1)$ and $b_{kn} = f(n-k+1)a_k$ for $k = 2, 3, \cdots, n-1$.

Since $a_1 = a_2 = 1$, it follows that $f(1) = f(2) = 1$; for $n \geqq 3$, we have

$$f(n) = 2f(n-1) + \sum_{k=2}^{n-1} a_k f(n-k+1) = 3f(n-1) + \sum_{k=2}^{n-2} f(k)f(n-k).$$

If we write $y = F(x) - x$, then it is not hard to see that this implies that $y = x^2 + 3xy + 2y^2$. Solving this for y, we obtain

$$F(x) = \tfrac{1}{4}\{1 + x - \sqrt{1 - 6x + x^2}\},$$

the negative sign being taken since $F(0) = 0$. But the Gegenbauer polynomials $C_n^\nu(z)$ have the generating function

$$(1 - 2zx + x^2)^{-\nu} = \sum_{n=0}^{\infty} C_n^\nu(z)x^n,$$

so that, for $n \geqq 2$, we have

$$f(n) = -\tfrac{1}{4}C_n^{-1/2}(3).$$

Using the asymptotic expression for the Gegenbauer polynomials as given in

G. Szegö, *Orthogonal Polynomials*, AMS Colloquium Publications, Vol. 23, 1959, pp. 194–195, we see that

$$f(n) = \frac{\sqrt{3\sqrt{2}-4}}{4\sqrt{\pi}} \frac{(3+2\sqrt{2})^n}{n\sqrt{n}} \left\{1 + \frac{3(8-3\sqrt{2})}{32n} + O(n^{-2})\right\}.$$

II. *Comment by D. E. Knuth, Stanford University.* The problem was originally posed and solved by Ernst Schröder as one of his famous "four combinatorial problems." (See Zeit. für Math. 15 (1870), 361–376.) I don't think that Schröder gave the asymptotic value, but it can be found in my book, *Fundamental Algorithms*, Addison-Wesley, 1968, pp. 534 and 587.

Also solved by D. A. Darling and by the proposer. Partial solutions by M. G. Greening (Australia), Harry Lass, and P. L. Montgomery.

A Totient Inequality

E 2316 [1971, 904]. *Proposed by R. S. Luthar, University of Wisconsin at Janesville*

Show that

$$\phi(n^2) + \phi(n^2+2n+1) < 2n^2,$$

where n is any integer > 2.

I. *Solution by Stephen Spindler, University of Chicago.* We have $\phi(n) \leqq n-1$ with equality if and only if n is prime. Thus

$$\begin{aligned}\phi(n^2) + \phi(n^2+2n+1) &= n\phi(n) + (n+1)\phi(n+1)\\ &\leqq n(n-1) + (n+1)n = 2n^2,\end{aligned}$$

with equality if and only if n and $n+1$ are both prime, i.e., if and only if $n = 2$.

II. *A sharper result by David Sumner, University of South Carolina, Columbia.* Indeed we have

$$\phi(n^2) + \phi(n^2+2n+1) \leqq \begin{cases}\frac{1}{2}(3n^2+2n) & \text{if } n \text{ is even,}\\ \frac{1}{2}(3n^2+1) & \text{if } n \text{ is odd.}\end{cases}$$

Proof. Suppose n is even. Then

$$n\phi(n) + (n+1)\phi(n+1) \leqq n(n/2) + (n+1)n = \tfrac{1}{2}(3n^2+2n).$$

Suppose n is odd. Then

$$n\phi(n) + (n+1)\phi(n+1) \leqq n(n-1) + (n+1)(n+1)/2 = \tfrac{1}{2}(3n^2+1).$$

III. *Generalization by Gerald Bergum, South Dakota State University.* We shall show that, for any integers $n > k \geqq 2$,

$$\phi(n^k) + \phi((n+1)^k) < kn^k.$$

It is easily established by mathematical induction that $(1-1/n) + (1+1/n)^{k-1} \leqq k$ for $n > k \geqq 2$. Since not both n and $n+1$ can be primes, we have

$$\begin{aligned}\phi(n^k) + \phi((n+1)^k) &= n^{k-1}\phi(n) + (n+1)^{k-1}\phi(n+1)\\ &< n^{k-1}(n-1) + (n+1)^{k-1}n = n^k((1-1/n) + (1+1/n)^{k-1}) \leqq kn^k.\end{aligned}$$

IV. *Generalization by David Zeitlin, Minneapolis.* We shall show that, for $n > 2$, $k \geqq 2$,

$$\phi(n^k) + \phi((n+1)^k) < 2n^2(n+1)^{k-2}.$$

This can be established by mathematical induction; the "induction step" is

$$\begin{aligned}\phi(n^{k+1}) + \phi((n+1)^{k+1}) &= n\phi(n^k) + (n+1)\phi((n+1)^k)\\ &< (n+1)[\phi(n^k) + \phi((n+1)^k)]\\ &< (n+1)2n^2(n+1)^{k-2} = 2n^2(n+1)^{k-1}.\end{aligned}$$

V. *Generalization by C. S. Venkataraman and R. Sivaramakrishnan, Trichur, India.* We shall show that, for $m, n > 2$,

$$\phi(mn) + \phi((m+1)(n+1)) < 2mn.$$

Assuming $m \geqq n$, we have

$$\begin{aligned}\phi(mn) + \phi((m+1)(n+1)) &\leqq m\phi(n) + (m+1)\phi(n+1)\\ &< m(n-1) + (m+1)n = 2mn - (m-n) \leqq 2mn.\end{aligned}$$

Also solved by 90 other readers.

A Totient Equation

E 2317 [1971, 905]. *Proposed by R. S. Luthar, University of Wisconsin at Janesville*

Find all pairs of natural numbers m, n such that

$$\phi(mn) = \phi(m) + \phi(n).$$

Solution by Irving Gerst, State University of New York at Stony Brook. Using the known relation $\phi(mn) = d\phi(m)\phi(n)/\phi(d)$, where $d = (m,n)$, we can write the given equation as $1/a + 1/b = d$, where $a = \phi(m)/\phi(d)$ and $b = \phi(n)/\phi(d)$. Since a and b are both positive integers, it follows that either $d = 2$ and $a = b = 1$, or

$d = 1$ and $a = b = 2$. The first case yields $\phi(m) = \phi(n) = 1$, whence $m = n = 2$, and the second case yields $\phi(m) = \phi(n) = 2$, giving one of m, n equal to 3 and the other equal to 4.

Also solved by the proposer and 74 other readers.

Lost in the Shuffle

E 2318 [1971, 905]. *Proposed by Thomas Hughes, Arlington, Texas*

Suppose that a machine is constructed to shuffle an ordinary 52-card deck in the same manner each time. How efficient could this machine be? That is, what is the maximum number of shuffles that could occur before the deck is returned to its original order?

Solution by C. V. Heuer, Concordia College. The question amounts to asking for the maximum order for an element in S_{52}, the symmetric group on 52 letters. Since the order of a permutation is the least common multiple of the lengths of the cycles which occur when the permutation is expressed as a product of disjoint cycles, we are looking for the maximum value of LCM(P), taken over all partitions P of 52. One can show that this maximum can be realized by using a partition $P = (x_1, \cdots, x_n)$ in which each x_i is unity or a prime power and where x_i and x_j are relatively prime for $i \neq j$. Checking these partitions yields the partition (1, 1, 1, 4, 5, 7, 9, 11, 13) as the cycle structure of an element of largest order, namely, 180,180.

For more information concerning the function $f(n) = \max\{o(x): x \in S_n\}$ see Jean-Louis Nicolas, *Sur l'ordre maximum d'un élément dans le groupe S_n des permutations*, Acta Arith. 14 (1968), 315–332 and C. V. Heuer, *Bounds on the least common multiple of integers with a fixed sum*, Dept. of Math. Preprints, No. 87, University of Oklahoma, Norman, Oklahoma.

Also solved by Edward Argyle, B. J. Bock, Bonnie Brusseau, B. R. Caine, Frederick Carty, C. S. Karuppan Chetty (India), M. S. Demos, R. L. Enison, Michael Goldberg, H. S. Hahn, C. P. McCarty, Eric Rosenthal, Michael Shimshoni (Israel), and G. J. Simmons.

At least a dozen readers were indeed lost in the shuffle — and contributed incorrect solutions.

Editor's comment. The problem is not new as several readers point out. It was first solved by W. H. H. Hudson in *Educational Times Reprints*, Vol. II (1865), p. 105. It is included in W. W. Rouse Ball, *Mathematical Recreations and Essays*, Revised Edition, 1963, 311–312; and it is mentioned by Martin Gardner in the November 1966 issue of *Scientific American* (p. 143). The mathematical counterpart to this problem is an exercise in F. E. Hohn, *Elementary Matrix Algebra*, Second Edition, 1964, p. 233. Argyle recommends an interesting discussion of shuffling in R. A. Epstein, *The Theory of Gambling and Statistical Logic*, Academic Press, 1967, Chap. 6.

Several solvers note that the addition of a joker raises the maximum length to 360,360 by allowing the cycle structure (5, 7, 8, 9, 11, 13). We note that in the 52-card deck, the cycle structures (1, 2, 4, 5, 7, 9, 11, 13) and (3, 4, 5, 7, 9, 11, 13) lead also to permutations of maximum order 180,180; the last might be considered a better shuffle since it leaves no card fixed.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before January 31, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5872*. *Proposed by Shmuel Schreiber, Bar-Ilan University, Israel*

Let C_n denote the region in Euclidean x space defined by $x_i \geqq 0$ for $i = 1, \cdots, n$ and $y_i \geqq 0$ for $i = 1, \cdots, n$, where

$$x_1 = 1 - 2y_1 + y_2$$

$$x_i = 1 + y_{i-1} - 2y_i + y_{i+1} \quad (2 \leqq i \leqq n-1)$$

$$x_n = 1 + y_{n-1} - 2y_n.$$

Prove that C_n is a convex polytope of the combinatorial type of a cube and that its volume is $(n+1)^{n-1}/(n!)$. (The result has some use in tournament theory.)

5873. *Proposed by Helge Tverberg, University of Bergen, Norway*

Those real polynomials in x and the greatest integer function $[x]$ which are continuous functions of x form a ring A, containing R. Find the minimal set of generators, over R, of A.

5874. *Proposed by T. E. Elsner, Michigan State University*

Let X be a compact T_0-space and let A be the set of closed singletons in X. Show that every subset containing A is compact.

5875. *Proposed by Anon, Erewhon-upon-Yarkon*

Suppose $f(t)$ is twice differentiable and

$$\lim_{t \to \infty} [f(t) + f'(t) + f''(t)] = L.$$

Prove $\lim_{t\to\infty} f(t) = L$. (Compare an exercise in Hardy, *Pure Mathematics*; $f + f' \to L$.)

5876.* *Proposed by C. H. Kimberling, University of Evansville*

In the ring of 2×2 matrices over the reals, is every unimodular matrix a product of matrices of finite order? If so, generalize.

5877. *Proposed by R. Shantaram, University of Michigan at Flint*

Let $\{a_n\}$ be a sequence of positive real numbers such that

$$\lim_{n\to\infty} (a_1 + a_2 + \cdots + a_n)/n = a, \quad 0 < a < \infty.$$

For $\alpha > 0$, find $\lim_{n\to\infty}(a_1^\alpha + a_2^\alpha + \cdots + a_n^\alpha)/n^\alpha$. What if $a = 0$?

SOLUTIONS OF ADVANCED PROBLEMS

The Spanning Trees of an n-Wheel

5795 [1971, 548]. *Proposed by B. R. Myers, University of Notre Dame*

An n-wheel is a graph consisting of one "outer" circuit having n vertices and edges along with the n edges connecting these vertices to a single "hub" vertex. A spanning tree of a graph on $(n + 1)$ vertices is a collection of n edges in the graph which contains no circuit.

How many different spanning trees are there in an n-wheel? (The result is conveniently expressible in terms of Fibonacci numbers.)

Solution by P. M. Gibson, University of Alabama in Huntsville. Define the $n \times n$ matrix A_n by letting

$$A_n = \begin{bmatrix} 3 & -1 & 0 \cdots & 0 & -1 \\ -1 & 3 & -1 \cdots & 0 & 0 \\ 0 & -1 & 3 \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & 3 & -1 \\ -1 & 0 & 0 \cdots & -1 & 3 \end{bmatrix}.$$

If the vertices of an n-wheel, $n \geqq 3$, are labeled so that $(v_1, v_2, \cdots, v_n, v_1)$ is the outer circuit, then by a theorem of Trent the number of spanning trees of this n-wheel is equal to the determinant of A_n [C. Berge, *The Theory of Graphs*, p. 159]. Let F_n be the nth Fibonacci number, and let B_n be the principal submatrix of A_n that remains after rows (and columns) 1 and 2 are removed. Expanding by the first row of B_{n+2} and simplifying, we have

$$\det B_{n+2} = 3 \det B_{n+1} - \det B_n. \tag{1}$$

Hence, since $\det B_3 = 3 = F_4$ and $\det B_4 = 8 = F_6$,

$$\det B_n = F_{2n-2}. \tag{2}$$

Expanding by the first row of A_n, simplifying, and then using (1) and (2) we obtain

$$\det A_n = 3\det B_{n+1} - 2\det B_n - 2 = F_{2n+2} - F_{2n-2} - 2.$$

Also solved by F. R. Bernhart, R. T. Bumby, P. J. Federico, Peter Hajdu, A. A. Jagtuin (Netherlands), D. J. Kleitman, E. C. Milner, C. C. Rousseau, K. Walker (England), Roger Weitzenkamp, M. R. Wise, and the proposer.

Editorial Notes. E. V. Milner points out that a solution is found in J. Sedlacek, *On the skeletons of a graph or digraph*, Proceedings of the Calgary International Conference of Combinatorial Structures and their Applications (1969), 387–391. The result given is

$$\left(\frac{3+\sqrt{5}}{2}\right)^n + \left(\frac{3-\sqrt{5}}{2}\right)^n - 2.$$

By direct analysis Bumby obtains the critical recursion equation $f_n - 3f_{n-1} + f_{n-2} = 0$ for the number of trees without reference to the matrix A_n. He starts with a graph formed by $(n+1)$ vertices $(0, 1, 2, \cdots, n)$ and segments (0,1), (0,2), $\cdots$, $(0,n)$, (1,2), (2,3), (3,4), $\cdots$, $(n\text{–}1,n)$, as pictured by a triangle and lines drawn from a vertex to $(n\text{–}2)$ division points on the opposite side.

The following note is offered by Federico with his solution: The matrix formula or theorem for the number of spanning trees in a graph was apparently first discovered by Brooks, Smith, Stone and Tutte from results in electrical network theory, some going back to Kirchhoff's paper of 1847, and appears in their classic paper of 1940 on *The dissection of rectangles into squares* (Duke Math. Journal 7 (1940), pp. 312 ff.). They refer to the number as the *complexity* of the graph and it plays a role in the theory of the dissection of rectangles. The theorem was rediscovered by Trent (H. M. Trent, *A note on the enumeration and listing of all possible trees in a connected linear graph*, Proc. Nat. Acad. Sci. 40 (1954), 1004–1007). An independent rediscovery is given by S. Okada and R. Onodera, *On network topology*, Bulletin of the Yamagata University, Natural Science, 2 (1952), (89–117). König has no reference to or suggestion of the theorem in his textbook of 1936 (*Theorie der endlichen und unendlichen Graphen*, Chelsea reprint, 1950), the first full dress text on graph theory, so presumably it was unknown at that time.

On Euler's Totient

5796 [1971, 549]. *Proposed by R. S. Luthar, University of Wisconsin, Janesville*

Show that, $\phi(n)$ being the Euler totient,

$$\limsup_{n\to\infty} \frac{\phi(n+1)}{\phi(n)} = \infty, \quad \liminf_{n\to\infty} \frac{\phi(n+1)}{\phi(n)} = 0.$$

Solution by Neal Felsinger, Edgewood Arsenal, Md. Let p_i be the ith prime, let r be a positive integer and $n = p_1p_2\cdots p_r$. By Dirichlet's theorem, for some k, $q = kn+1$ is prime. Then $\phi(q) = kn$ while

$$\phi(q-1) = kn \prod_{p|kn} (1-1/p) \leqq kn \prod_{i=1}^{r} (1-1/p_i).$$

Hence $\phi(q)/\phi(q-1) \geqq 1/\prod_{i=1}^{r}(1-1/p_i)$. Now it is well known that $\prod_{p\text{ prime}}(1-1/p)$ diverges to 0. Thus, letting r become large we have $\limsup \phi(n)/\phi(n-1) = \infty$.

For the second part, for some k, $q = kn-1$ is prime. Then $\phi(q) = kn - 2$ while

$$\phi(q + 1) = kn \prod_{p|kn} (1 - 1/p) \leqq kn \prod_{i=1}^{r} (1 - 1/p_i).$$

Thus $\phi(q + 1)/\phi(q) \leqq (kn/(kn - 2)) \prod_{i=1}^{r}(1 - 1/p_i)$. Letting r become large, we have $\liminf \phi(n + 1)/\phi(n) = 0$.

Also solved by D. W. Ballew, Paul Bateman, S. J. Benkoski, D. Borwein, Robert Breusch, R. T. Bumby, W. F. de la Vega (France), R. J. Dickson, R. E. Dressler, Leon Gerber, Robert Giese, Emil Grosswald, J. L. Hlavka, Vaclac Konecňy, E. S. Langford, Marijo LeVan, O. P. Lossers (Netherlands), Arthur Marshall, L. E. Mattics, P. L. Montgomery, Ivan Niven, Andrew Odlyzko, Bob Prielipp, C. A. Rofer, T. Šalát (Czechoslovakia), H. N. Shapiro, Allen Stenger, Karl Stoop (Colombia), D. Suryanarayana, J. H. van Lint, C. S. Venkataraman (India), and Konrad Victor (Israel).

Several solvers note that the result may be found in B. S. K. R. Somayajulu, *On Euler's Totient Function*, Math. Student 18 (1950), 31–32. Prielipp refers to the stronger result: $\phi(n + 1)/\phi(n)$, $n = 1, 2, \ldots$, is dense in the set of nonnegative real numbers (see Sierpinski, *Elementary Theory of Numbers*, Hafner, New York, 1964, p. 235–236). Several solvers note that the quotient may be replaced by $\phi(n + \alpha)/\phi(n)$. Further generalizations are cited by Andrzej Makowski: see A. Schinzel in Bull. Acad. Polon, Sci. Classe Troisième, 3 (1955) p. 415 ff. also 2 (1954), p. 463 ff. Grosswald and Shapiro establish

$$\limsup \frac{\phi(n + 1)}{\phi(n)\log\log n} > 0, \quad \liminf \frac{\phi(n + 1)\log\log n}{\phi(n)} < \infty.$$

Inverses in Prime Rings

5797 [1971, 549]. *Proposed by I. N. Herstein, University of Chicago, and Susan Montgomery, University of Southern California*

A theorem of Marshall Osborn states: *If R is a simple ring of characteristic not 2 and with an involution such that every nonzero symmetric element is invertible, then either R is a division ring or is 4-dimensional over its center.* Show that if R is a prime ring with involution, of characteristic 2, and if every nonzero symmetric element of R is invertible, then R must be a division ring. (Prime means $xRy = 0$ implies $x = 0$ or $y = 0$.)

Solution by G. J. Janusz, University of Illinois. Let x^* denote the image of x under the involution. In place of the full prime condition, we need use only $xRx^* = 0$ implies $x = 0$.

Suppose we know that $x \neq 0$ implies xRx^* contains a nonzero symmetric element xyx^*. Then it has an inverse u and so yx^*u is a right inverse for x. Thus every nonzero element has a right inverse; in particular x^* has a right inverse w and it follows that w^* is a left inverse for x. Thus every nonzero element has an inverse.

Now we must prove the supposition. Suppose 0 is the only symmetric element in xRx^*. For any r in R, $x(r+r^*)x^*$ is symmetric and hence is 0. This means

$$xrx^* = -xr^*x^* = -(xrx^*)^*.$$

By assumption, the characteristic is 2 so xRx^* is symmetric and thus equals 0. Thus $xRx^* = 0$ and $x = 0$ as we wished to prove.

Also solved by Cecilia H. Brook, D. Ž. Djoković, T. S. Erickson, A. A. Jagers (Netherlands), Gerald Losey, W. Margolis, Nadine C. Myers, E. J. Taft, and the proposers.

NOTE. Inexplicably this problem has been introduced as 5837 [1972, 94]. 5837 has been withdrawn and solvers of 5837 are included above.

Convex Properties of the Γ-function

5798 [1971, 549]. *Proposed by C. J. Eliezer, La Trobe University, Bundoora, Australia*

Prove that for $x > 1$ and $y > 1$,

$$\frac{\Gamma(x)}{(x-1)^x} + \frac{\Gamma(y)}{(y-1)^y} \geqq \frac{2\Gamma\left(\frac{x+y}{2}\right)}{\left(\frac{x+y}{2}-1\right)^{(x+y)/2}},$$

$$\frac{\Gamma(x)\Gamma(y)}{\left[\Gamma\left(\frac{x+y}{2}\right)\right]^2} \geqq \frac{(x-1)^x(y-1)^y}{\left(\frac{x+y}{2}-1\right)^{x+y}}.$$

Solution by Myron Lipow, Palos Verdes Penin, California. Relying on the well-known Laplace Transform formula:

$$\Gamma(z) = s^z \int_0^\infty e^{-st}t^{z-1}dt$$

valid for $\mathrm{Re}(z) > 0$ and $\mathrm{Re}(s) > 0$, hence if $\mathrm{Re}(z) > 1$ and if $s = z - 1$ then

$$\frac{\Gamma(z)}{(z-1)^z} = \int_0^\infty (te^{-t})^{z-1}dt.$$

Thus, for real x, $y > 1$,

$$\text{(1)} \qquad \frac{\Gamma(x)}{(x-1)^x} + \frac{\Gamma(y)}{(y-1)^y} = \int_0^\infty ((te^{-t})^{x-1} + (te^{-t})^{y-1})\, dt.$$

Since $(a^{(x-1)/2} - a^{(y-1)/2})^2 \geqq 0$ for $a \geqq 0$, we have

$$a^{x-1} + a^{y-1} \geqq 2a^{(x+y)/2-1}.$$

Hence the right-hand side of (1) is

$$\geqq 2\int_0^\infty (te^{-t})^{(x+y)/2-1}dt$$
$$= 2\Gamma\left(\frac{x+y}{2}\right)\Big/\left(\frac{x+y}{2}-1\right)^{(x+y)/2},$$

using the previously stated Laplace Transform. This proves the first proposed inequality.

For the second inequality we have

$$(2)\qquad \frac{\Gamma(x)}{(x-1)^x}\,\frac{\Gamma(y)}{(y-1)^y} = \int_0^\infty (te^{-t})^{x-1}dt\int_0^\infty (te^{-t})^{y-1}dt$$
$$\geqq \left(\int_0^\infty (te^{-t})^{(x-1)/2}(te^{-t})^{(y-1)/2}dt\right)^2$$

by Schwarz' inequality. The right-hand side of (2)

$$= \left(\Gamma\left(\frac{x+y}{2}\right)\right)^2\Big/\left(\frac{x+y}{2}-1\right)^{x+y}.$$

Upon rearranging, we get the desired result.

Also solved by R. J. Dickson, M. G. Greening (Australia), S. A. Greenspan, A. A. Jagers (Netherlands), Hans Kappus (Germany), Beatriz Margolis (Argentina), I. Olkin, Yi-Chuan Pan, P. G. Rooney, David Shelupsky, F. W. Steutel (Netherlands), Brian Thorpe, and the proposer.

NOTE. Most of the solvers established by differentiation the convexity of the function $\Gamma(x)/(x-1)^x$ and its logarithm, and thence obtained the results. The misprint in the first of the proposed inequalities (corrected above) was noted by all solvers.

Lower Bounds for an Alternating Series

5799 [1971, 549]. *Proposed by C. J. Eliezer, La Trobe University, Australia*

Prove that for $-1 < p < 1$,

$$\frac{1}{p+1}-\frac{1}{p+2}+\frac{1}{p+3}-\cdots \geqq \frac{1-4p+2p^2}{(1-p)(2-p)},$$

and

$$\frac{1}{p+1}-\frac{1}{p+2}+\frac{1}{p+3}-\cdots \geqq \frac{(1-p)(2-p)}{(3-2p)}.$$

Solution by Edward Severn, Undergraduate, Cedarbrae Collegiate Institute, Scarborough, Ontario. From $a+1/a>2$, $a\neq 1$, $a>0$, we have

$$\frac{t^p}{t+1}+\frac{t+1}{t^p}>2,\quad 0<t<1,\quad -1<p<1,$$

$$\int_0^1 \frac{t^p}{t+1}dt > 2 - \int_0^1 \frac{t+1}{t^p}dt,$$

whence by immediate calculation we have

$$\frac{1}{p+1} - \frac{1}{p+2} + \frac{1}{p+3} - \cdots > 2 - \frac{1}{2-p} - \frac{1}{1-p} = \frac{1-4p+2p^2}{(1-p)(2-p)}.$$

We have also the inequality (Schwarz)

$$\int_a^b f(x)dx \cdot \int_a^b \frac{dx}{f(x)} \geqq (b-a)^2.$$

If we set $a = 0$, $b = 1$, $f(x) = x^p/(x+1)$, it follows that

$$\frac{1}{p+1} - \frac{1}{p+2} + \frac{1}{p+3} - \cdots = \int_0^1 \frac{x^p dx}{x+1} > \left[\int_0^1 \frac{x+1}{x^p}dx\right]^{-1}$$
$$= \frac{(1-p)(2-p)}{3-2p}.$$

As a matter of fact, the first part of the problem follows from the second upon setting $a = (1-p)(2-p)/(3-2p)$ in the inequality $a + 1/a > 2$.

Also solved by D. Borwein, M. G. Greening (Australia), B. H. Harris, A. A. Jagers (Netherlands), R. E. Shafer, L. E. Ward, Sr., and the proposer.

Locally Null Subsets of a Locally Compact Group

5800 [1971, 549]. *Proposed by Joel Pitcairn, Huntingdon Valley, Pa.*

Exercise 16.1 of Halmos, *Measure Theory* says: If E is a Lebesgue measurable set such that, for every x in a dense set, $\mu(E\Delta(E+x)) = 0$, then $\mu(E) = 0$ or $\mu(E') = 0$. Prove the following generalization (which is useful for producing 'maximally non-measurable' sets): If E is a subset of a locally compact group (with left Haar measure μ) such that, for every x in a dense set, $E\Delta xE$ is locally null, then either (1)E is locally null or (2) E' is locally null or (3) $\mu^*(A \cap E) = \mu^*(A \cap E') = \mu(A)$ for every measurable set A. (A set is locally null if its intersection with every compact set has measure 0.)

Solution by the proposer. The set function defined by $\lambda(A) = \mu^*(A \cap E)$ is a Borel measure (countable additivity follows from Halmos 11.B). If A is a Borel set of finite measure, then for all x and y we have

$$\lambda(xA) = \lambda(xA \cap yA) + \lambda(xA \cap yA') \leqq \lambda(yA) + \mu(xA\Delta yA).$$

Interchanging x and y yields $|\lambda(xA) - \lambda(y(A)| \leqq \mu(xA\Delta yA)$. Now $\mu(xA\,\Delta\, yA) \to 0$ as $x^{-1}y \to e$ (Halmos, 61. A), and so $\lambda(xA)$ is a continuous function of x.

If $E\Delta xE$ is locally null, $\lambda(A) = \mu^*(A \cap E) = \mu^*(A \cap xE) = \mu^*(x^{-1}A \cap E) = \lambda(x^{-1}A)$. Since this holds on a dense set, continuity implies that it holds for all x, i.e., that A is left-invariant under λ. But every Borel set is a limit of an increasing sequence of Borel Sets of finite measure, and it follows that every Borel set is left-invariant under λ. This shows that λ is a left Haar measure (or zero), and so by uniqueness there exists α such that for every measurable set A, $\mu^*(A \cap E) = \alpha\mu(A)$; and $0 \leqq \alpha \leqq 1$ since $\lambda \leqq \mu$. Now $E\Delta xE = E'\Delta xE'$ and the same argument gives us β, $0 \leqq \beta \leqq 1$, such that $\mu^*(A \cap E') = \beta\mu(A)$ for every measurable set A. If B has finite outer measure, then for every open set U containing B, $\mu^*(B \cap E) \leqq \mu^*(U \cap E) = \alpha\mu(U)$; hence $\mu^*(B \cap E) \leqq \alpha\mu^*(B \cap E)$. So if $\alpha < 1$, E is locally null. Similarly, if $\beta < 1$, E' is locally null. The only other possibility is $\alpha = \beta = 1$, in which case (3) holds.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR., AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges.

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057. All unsigned material is written by one of the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor to avoid duplication.*

C *Set Theory and Topology*. By Philip Nanzetta and George E. Strecker. Bogden & Quigley, New York, 1971. ix + 117 pp. $8.50. (Telegraphic Review, November 1971.)

This book is designed for a course taught by the R. L. Moore method which assumes, among other things, that all proofs are presented by the students while the instructor acts as moderator, referee, and sometimes cheerleader, but, most important of all, keeps his proofs to himself. Nanzetta and Strecker give at most two or three proofs (and these are quite elementary, serving to set the style of rigor) and give only mild hints as to how to prove the more difficult theorems. So this book, unlike all other textbooks in topology (a contradiction in terms to a Moore man) will neither contaminate the mind nor harm the creative potential of the beginning student of topology who uses it.

I used the book in a one semester graduate course in Introductory Topology which is offered in a masters degree program where no PhD program exists. Even though my students lacked some of the drive and competitive instinct characteristic of PhD candidates, I found the Moore method and the use of Nanzetta–Strecker far superior to the usual lecture course.

While the authors "have included no category theory in the text, seeds of category

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 **NUMBER 9**

CONTENTS

(Continued on inside cover)

NOVEMBER **1972**

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

THE GEOMETRY OF RADON'S THEOREM

B. B. PETERSON, Middlebury College and the University of Washington

1. Introduction. The closely related theorems of Helly, Caratheodory, and Radon form a crucial triumvirate in the theory of convex bodies. All are familiar, but that of Radon seems to have received somewhat less attention than its partners. Our intention here is to rectify that situation by geometric considerations which lead clearly and intuitively to Radon's result. The general procedure will be to find two d-simplices which meet in a common $(d-1)$-face and arrange an induction based on the hyperplane carrying that face. This double-simplex configuration can be exploited to establish each of the three theorems as well as many extensions and generalizations.

RADON'S THEOREM. *If T is a set of k points ($k \geqq d+2$) in Euclidean d-space, there are disjoint sets T_1 and T_2 with $T = T_1 \cup T_2$ and $conv\, T_1 \cap conv\, T_2 \neq \emptyset$.*

The briefest proofs of this theorem and its relatives are algebraic in nature and among the most elegant of the genre. As is often the case, however, the elegance, compactness and power of these methods may tend to mask the underlying geometric situation. (Of course this is advantage as well as problem, and the occurrence is by no means peculiar to Algebra and Geometry. Whether the gains in one area outweigh the costs in another is a matter of personal prejudice.) Our proof, which is essentially that of Rado [**12**], depends only on the fact that it is in general possible to separate some two points from a collection of $d+2$ by a hyperplane containing the remaining d, or, equivalently, to find in the collection the vertices of the double simplex configuration. Several results characterizing the sets T_1 and T_2 and extending the main theorem seem to follow more intuitively and, in some cases, more easily from this separation property than from their more familiar algebraic settings. Both Helly's Theorem and Steinitz's Theorem on the interior of convex hulls are easily derived from the same central configuration. Some of the proofs may be more easily followed in 3-space where the geometry is most clearly revealed.

2. Notations and definitions. All results and definitions are stated for Euclidean d-space E^d although the proofs are in most cases valid for a vector space over an arbitrary ordered field. Certain sets determined by two points x and y will be denoted as follows: closed segment by $[xy]$, open segment by (xy), half-open segments by $(xy]$ and $[xy)$, closed and open rays from x containing y by $[xy$ and $(xy$, line by xy. The set consisting of x alone is denoted by $\{x\}$. An ***n*-set** is a set containing exactly n points.

Bruce Peterson received his Syracuse Ph.D. in 1962 under E. Hemmingsen. He has been at the Middlebury College since then where he is now a Professor and Chairman of the Department. His main research interests are convexity and particularly sets of constant width. He spent 1970–1971 as a Visitor at the University of Washington. *Editor.*

A set is **convex** if it contains the closed segment joining any two of its points. A **k-flat** is a k-dimensional affine subspace of E^d (a translate of a k-dimensional linear subspace of E^d). For any set T the **convex hull** of T, denoted by conv T, is the intersection of all convex sets containing T; it is the minimal convex set containing T. Similarly, the **affine hull**, aff T, is the intersection of all flats containing T; it is the minimal flat containing T. A **hyperplane** is a $(d-1)$-flat. At each point x of the boundary of a convex set S there is a hyperplane, called a **supporting hyperplane**, which contains x and is so situated that one of the two closed half-spaces it determines entirely contains S. A **polytope** is the convex hull of a finite set. A polytope P is a **k-polytope** if aff P is a k-flat. If $P = \operatorname{conv} T$ is a polytope, the minimal subset T' of T for which $P = \operatorname{conv} T'$ is called the **vertex set** of T and denoted vert P.

A finite set S is in **general position** if, for $k < d$, no $(k+2)$-subset of S lies in a k-flat. A **k-simplex,** $k \leqq d$, is the convex hull of a $(k+1)$-set in general position. If σ is a k-simplex and $T' \subset \operatorname{vert} \sigma$ is an n-set, the set $\operatorname{vert} \sigma - T'$ determines a $(k-n)$-simplex called a **$(k-n)$-face of** σ. A **facet** is a $(k-1)$-face. Note that the affine hull of a k-simplex is always a k-flat.

The pair $\{T_1, T_2\}$ is a **partition** of the set T if $T_1 \neq \emptyset$, $T_2 \neq \emptyset$, $T = T_1 \cup T_2$, and $T_1 \cap T_2 = \emptyset$. The partition is a **Radon partition** if, in addition, $\operatorname{conv} T_1 \cap \operatorname{conv} T_2 \neq \emptyset$. A set S is **(r,k)-divisible** if there are r disjoint non-empty sets whose union is S and whose convex hulls intersect in a set of dimension at least k.

The following result should be unsurprising; a proof may be found in Grünbaum [**5**, p. 31]. The corollary will be used to establish the Separation Lemmas of sections 3 and 5.

THEOREM. *Any d-polytope P in E^d is the intersection of the closed halfspaces containing P and determined by the facets of P.*

COROLLARY. *If P is a polytope and $x \notin P$, then some facet F of P determines a supporting hyperplane aff F which separates x from P.*

Proof. Since $x \notin P$, it is not contained in some closed halfspace containing P and determined by a facet of P.

3. A Proof of Radon's Theorem. We first establish a separation lemma. Alternate proofs of the lemma are not difficult to find, although, with the exception of Rado's paper, the author knows of no place where they appear explicitly. While simple and intuitive, the statement seems worthy of mention if only for the ubiquity of the configuration it assures.

SEPARATION LEMMA 1. *If $T = \{v_0, v_1, \cdots, v_{d+1}\}$ is a $(d+2)$-subset of E^d and does not lie in a hyperplane, then there is a hyperplane on a d-subset of T which separates the remaining two points.*

Proof. Since the statement is trivially true for $d = 1$, we may proceed by in-

duction on the dimension d, assuming the lemma proved for $(d+1)$-sets in E^{d-1}. If $d+1$ of the points v_μ lie in a hyperplane π (assume $v_0 \notin \pi$), two of these may be separated in π by a $(d-2)$-flat π' on the remainder. These two are separated in E^d by the hyperplane $\operatorname{aff}(\pi' \cup \{v_0\})$.

Otherwise, let $P = \operatorname{conv} T$ and assume v_0 is a vertex of P. Then $\operatorname{conv}(T - \{v_0\})$ is a d-simplex σ, and $v_0 \notin \sigma$. But then, by the corollary of section 2, some supporting hyperplane determined by a facet of σ separates v_0 from σ. This is the desired hyperplane.

Proof (*Radon's Theorem*). We shall proceed by induction on the dimension d, first finding a separating hyperplane and then projecting the remaining points onto it. The statement is obvious for $d = 1$; we assume it for $(d+1)$-sets in E^{d-1}. Let $T = \{v_0, v_1, v_2, \cdots, v_k\}$ be a k-subset of E^d. If any $(d+1)$-subset of T lies in a hyperplane, it admits a Radon partition. Adjoining the remaining points to either set in that partition will result in a Radon partition of T. By the same reasoning the case $k = d+2$ is sufficient to establish the theorem for all $k \geqq d+2$.

Assume then that $k = d+2$ and that no $(d+1)$-subset of T lies in a hyperplane. Let $\sigma_i = \operatorname{conv}(T - \{v_i\})$ for $i = 0, 1, 2, \cdots, d+1$ and $\sigma_{ij} = \operatorname{conv}(T - \{v_i\} - \{v_j\})$ for $i, j = 0, 1, 2, \cdots, d+1$. Applying the separation lemma and adjusting our subscripts if necessary, we may assume that v_0 and v_1 are separated by the hyperplane $\pi = \operatorname{aff} \sigma_{01}$ (see figure 1).

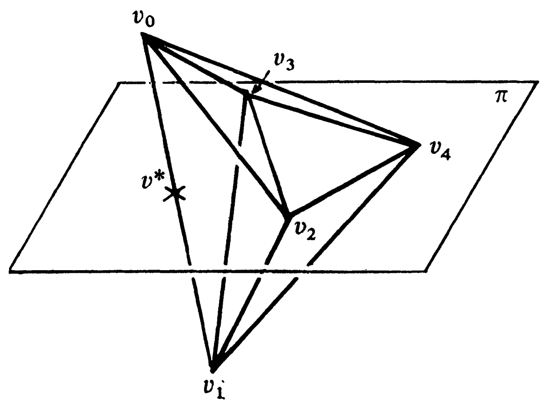


FIG. 1

Let $v^* = [v_0 v_1] \cap \pi$. If $v^* = v_j$ for some $j \neq 0$, $j \neq 1$, then $T_1 = \{v_j\}$ and $T_2 = T - \{v_j\}$ is the desired Radon partition.

Otherwise we consider the set

$$S = \{v^*, v_2, \cdots, v_{d+1}\}.$$

By the induction hypothesis there is a Radon partition $\{S_1, S_2\}$ of S. Assuming that $v^* \in S_1$ we take

$$T_1 = (S_1 - \{v^*\}) \cup \{v_0\} \cup \{v_1\} \text{ and } T_2 = S_2.$$

That $\{T_1, T_2\}$ is a partition of T is clear. From the fact that $v^* \in [v_0 v_1]$ it follows that $\operatorname{conv} S_1 \subset \operatorname{conv} T_1$ and $\operatorname{conv} T_1 \cap \operatorname{conv} T_2 \supset \operatorname{conv} S_1 \cap \operatorname{conv} S_2 \neq \emptyset$. This completes the proof.

4. Properties of the partition. What can be said about the sets T_1 and T_2? Without specifying that the points be in general position, very little. With that additional hypothesis, however, considerably more is known. We present the following results as corollaries because they follow so directly from geometric configuration of Separation Lemma 1.

COROLLARY 4.1. *Let T be a $(d+2)$-set in E^d. Then T is in general position if and only if the partition $\{T_1, T_2\}$ guaranteed by Radon's Theorem is unique.*

Proof. Adopting the same notation as in the proof of Radon's Theorem, we shall consider the mapping

$$\phi : \sigma_0 \to \pi$$

defined by $\phi(x) = xv_0 \cap \pi$. For any $\mu \neq 0$, $\mu \neq 1$ we have $v_\mu \in \sigma_{01}$ and $\phi(v_\mu) = v_\mu$. Clearly $\phi(v_1) = v^*$.

We first show that if T is in general position, then the partition is unique. We proceed by induction noting that the statement is trivial for $d = 1$. We assume it for $d - 1$. If a $(k+1)$-subset K of S lies in a $(k-1)$-flat, then $\phi^{-1}(K) \cup \{v_0\}$ is a $(k+2)$-set in a k-flat. Hence, if T is in general position so is S, and we may assume the partition $\{S_1, S_2\}$ is unique.

For any Radon partition $\{U_1, U_2\}$ of T, let $\{U_1', U_2'\}$ be the partition of $T - \{v_0\}$ given by $v_\mu \in U_k'$ if $v_\mu \in U_k$. Assume that $v_1 \in U_1'$ so that $U_1' \neq \emptyset$. If $U_2' = \emptyset$, then $U_2 = \{v_0\}$ and, since $\{U_1, U_2\}$ is a Radon partition,

$$v_0 \in \operatorname{conv} U_1 = \sigma_0.$$

Since this is impossible, $U_2' \neq \emptyset$. Obviously $U_1' \cap U_2' = \emptyset$, $\phi(U_1') \neq \emptyset$ and $\phi(U_2') \neq \emptyset$. Since v_1 is the only v_μ moved by ϕ, if $\phi(U_1') \cap \phi(U_2') \neq \emptyset$, we must have

$$\phi(v_1) \in \phi(U_2') = U_2' \subset U_2.$$

This is a contradiction since, by the general position hypothesis, $\phi(v_1) \neq v_\mu$ for $\mu \neq 1$. This establishes that $\{\phi(U_1'), \phi(U_2')\}$ is a partition of S.

Now let $z \in \operatorname{conv} U_1 \cap \operatorname{conv} U_2$ and distinguish two cases depending upon the placement of v_0.

Case 1: $v_0 \in U_1$. Then $U_2' = U_2 \subset \sigma_{01}$ and $\phi(U_2') = U_2'$. Hence

$$z \in \operatorname{conv} U_2 = \operatorname{conv} U_2' = \operatorname{conv} \phi(U_2').$$

Since $\phi(v_1) \in [v_0 v_1] \subset \operatorname{conv} U_1$ and $z \in \sigma_{01}$, we have

$$z = \phi(z) \in (\operatorname{conv} U_1) \cap \sigma_{01} = \operatorname{conv} \phi(U_1').$$

Case 2: $v_0 \in U_2$. Then $U_1' = U_1 \subset \sigma_0$ and $z \in \operatorname{conv} U_1 \subset \sigma_0$. Since $z \in \operatorname{conv} U_2 \subset \sigma_1$, we have $z \in \sigma_{01}$. Hence

$$z = \phi(z) \in \operatorname{conv} U_2' = \operatorname{conv} \phi(U_2').$$

Moreover

$$z = \phi(z) \in \operatorname{conv} [(U_1 - \{v_1\}) \cup \{\phi(v_1)\}] = \operatorname{conv} \phi(U_1').$$

Thus, in either case, $\{\phi(U_1'), \phi(U_2')\}$ is a Radon partition of S with $\phi(v_1) \in \phi(U_1')$. Since the partition $\{S_1, S_2\}$ is unique we must have

$$\phi(U_1') = S_1 \text{ and } \phi(U_2') = S_2.$$

In other words, any two Radon partitions of T can differ only in the placement of v_0. Reversing the roles of v_0 and v_1 and mapping σ_1 into π, the same two partitions can differ in the placement of v_1 only. This establishes the uniqueness of the partition $\{T_1, T_2\}$.

The converse is simple. If T is not in general position, some $(k+2)$-subset lies in a k-flat for $k < d$. There is a Radon partition of that subset, and, since the remaining points of T may be distributed arbitrarily without destroying "Radonness," the original partition is not unique.

COROLLARY 4.2. *Let $\{T_1, T_2\}$ be the Radon partition of a $(d+2)$-set in general position, then $\operatorname{conv} T_1 \cap \operatorname{conv} T_2$ is a single point.*

Proof. We proceed again by induction on d. Our induction hypothesis assures that $\operatorname{conv} S_1 \cap \operatorname{conv} S_2$ is the single point z. From the proof of Corollary 4.1 we know that

$$\operatorname{conv} T_1 \cap \operatorname{conv} T_2 \subset \sigma_{01}.$$

Hence $w \in \operatorname{conv} T_1 \cap \operatorname{conv} T_2$ implies that $\phi(w) = w$ and

$$w \in \operatorname{conv} \phi(T_1') \cap \operatorname{conv} \phi(T_2') = \operatorname{conv} S_1 \cap \operatorname{conv} S_2 = z.$$

Thus $\operatorname{conv} T_1 \cap \operatorname{conv} T_2$ consists of the point z alone.

Proskuryakov [**11**] and Kosmak [**10**] have shown that if T is in general position, then two points belong to the same member of the Radon partition if and only if they are separated by the hyperplane on the remainder. The geometric situation actually appears to suggest a slightly stronger form of the same result.

COROLLARY 4.3. *If the $(d+2)$-set T is in general position, then the point $\operatorname{conv} T_1 \cap \operatorname{conv} T_2$ is the intersection of all hyperplanes determined by d-subsets of T and separating the remaining two points of T.*

Proof. The partition $\{T_1, T_2\}$ is unique, but any hyperplane on a d-subset of T

and separating the remaining two points could have been used to prove the theorem. Obviously then z belongs to every such plane.

We prove that z is the only such point by induction on d, assuming the statement proved for $d-1$. The point z belongs to the hyperplane π and is the intersection of all $(d-2)$-flats on $(d-1)$-subsets of S and separating the remaining two points of S. Since $\operatorname{conv} S_1 \cap \operatorname{conv} S_2$ is the single point z, there must be at least $d-1$ such flats. Call these $(d-2)$-flats π'_μ for $\mu = 1,2,\cdots,d-1$. Consider the d hyperplanes $\pi_0 = \pi$ and $\pi_\mu = \operatorname{aff}(\pi'_\mu \cup \{v_0\})$ for $\mu = 1,2,\cdots,d-1$.

If $\phi(v_1) \in \pi'_j$, both v_0 and v_1 lie in π_j and the points of S separated by π'_j are points of T separated by π_j. The flat π'_j contains a $(d-1)$-subset) of Sa nd a $(d-2)$-subset of T, so that π_j contains a d-subset of T. Otherwise $\phi(v_1$ and some $v_k \in S$ are separated in π by the $(d-2)$-flat π'_j. Therefore v_1 and v_k are separated in E^d by π_j. In this case π'_j already contains a $(d-1)$-subset of T, so that π_j contains a d-subset of T.

All the flats π'_μ contain z and $\bigcap_{\mu=1}^{d-1} \pi_\mu$ is the line $v_0 z$. Hence $\bigcap_{\mu=0}^{d-1} \pi_\mu = z$. But then the intersection of all hyperplanes of the hypothesis is z and the theorem is proved.

We can count the separating hyperplanes on d-subsets of T rather simply. Assuming T_1 is a k-set, where $1 \leqq k \leqq [d/2]+1$ ($[x]$ being the greatest integer $\leqq x$), the total number of such hyperplanes is

$$\binom{k}{2} + \binom{d+2-k}{2},$$

which as a function of k is decreasing and bounded below by d in the indicated domain. Hence, we cannot assume from the general position of T that the flats determined by T are in "general position." More precisely, we can expect families of more than k hyperplanes to intersect in a $(d-k)$-flat (e.g., in E^3 we may well have four planes determined by T intersecting in a point). This difficulty complicates the generalization to (r,k)-divisibility. Reay **[14]** has shown, under the more stringent hypothesis of strong general position, which in effect rules out such bothersome intersections, that any $[(d+1)(r-1)+k+1]$-set in E^d is (r,k)-divisible.

A known result from the theory of convex polytopes **[5]** provides an interesting sidelight. Since the total number of hyperplanes on d-subsets of T is $\binom{d+2}{2}$, the number of these which support $\operatorname{conv} T$ is

$$\binom{d+2}{2} - \binom{k}{2} - \binom{d+2-k}{2} = (d+2-k)k.$$

Hence the simplicial polytope $\operatorname{conv} T$ has $(d+2-k)\,k$ facets, where $1 \leqq k \leqq [d/2] + 1$. By slightly pushing vertices any simplicial polytope may be considered as the convex hull of a collection of points in general position. Hence, this number yields

minimum and maximum values for the number of facets of a simplicial d-polytope with $d + 2$ vertices. (Note that the minimum occurs for $k = 2$, since, if $k = 1$, T_1 is a single point and $\operatorname{conv} T = \operatorname{conv} T_2$ has only $d + 1$ vertices.)

Kenelly and Hare [**8**] proved the following surprising characterization of Radon partitions by algebraic methods. That spheres should enter the picture at all seems strange although the proof requires little more than the separation lemma and a few elementary facts about spheres. In addition to the previous notation, we define S_μ to be the $(d-1)$-sphere on $\operatorname{vert}\sigma_\mu$ for $\mu = 0$ and 1. The bounded component of $E^d - S_\mu$ is denoted by $\operatorname{int} S_\mu$ and the d-ball $S_\mu \cup \operatorname{int} S_\mu$ by B_μ.

COROLLARY 4.4. *If the $(d + 2)$-set T is in general position and does not lie on a $(d-1)$-sphere, then v_0 and v_1 belong to the same member of the Radon partition of T if and only if*

(i) $v_1 \in \mathit{int}\, S_1$ *implies* $v_0 \in \mathit{int}\, S_0$ *and*

(ii) $v_1 \in E^d - B_1$ *implies* $v_0 \in E^d - B_0$.

That is, each point lies inside the sphere determined by the remaining $(d + 1)$-set, or each lies outside.

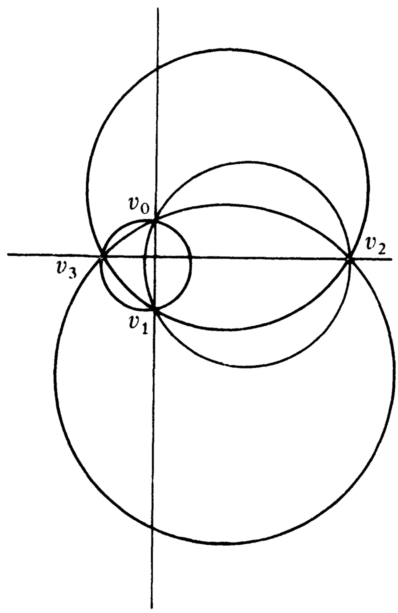


FIG. 2

(The situation for E^2 is pictured in Figure 2. Note that v_0 and v_1 satisfy (i) while v_2 and v_3 satisfy (ii).)

Proof. The spheres S_0 and S_1 intersect in a $(d-2)$-sphere which contains $\operatorname{vert}\sigma_{01}$. Together with v_0 (or v_1) $\operatorname{vert}\sigma_{01}$ determines S_1 (or S_0). Let H_μ, for $\mu = 0$ and 1,

be the closed halfspace determined by the hyperplane π and containing v_μ. Then π divides the balls B_0 and B_1 so that for each μ

$$B_0 \cap H_\mu \subset B_1 \quad \text{or} \quad B_1 \cap H_\mu \subset B_0$$

and

$$B_0 \cap H_1 \subset B_1 \text{ if and only if } B_1 \cap H_0 \subset B_0.$$

To prove (i) note first that $v_1 \in \operatorname{vert}\sigma_0 \subset S_0$. If $v_1 \in \operatorname{int} S_1$, there is a neighborhood $N(v_1)$ contained in $B_1 \cap H_1$ and meeting $E^d - B_0$. In particular

$$B_1 - (B_0 \cap H_1) \neq \varnothing.$$

Therefore $B_0 \cap H_1 \subset B_1$ and $B_1 \cap H_0 \subset B_0$. But v_0, which is not on S_0, lies in $B_1 \cap H_0 \subset B_0$ and must therefore belong to $\operatorname{int} S_0$. The proof is completed by reversing the roles of v_0 and v_1.

5. Extensions and generalizations. As mentioned previously, a powerful extension of Radon's theorem has been proved by Reay under a more restrictive dispersion hypothesis. The question of (r, k)-divisibility of sets which are merely in general position remains unsettled. We shall concern ourselves here only with the extension to $(2, k)$-divisibility of sets in general position, also solved by Reay [**14**]. Our approach via a stronger separation lemma is suggested by the observation that Separation Lemma 1 and Corollary 4.3 actually provide a separating hyperplane on a d-subset of T for every point not contained in the convex hull of the remainder.

SEPARATION LEMMA 2. *If T is a $(d+k)$-set in general position in E^d and if $v_0 \in T - conv(T - \{v_0\})$, then there is a hyperplane on a d-subset of T which supports $conv(T - \{v_0\})$ and separates it from v_0.*

Proof. By the corollary in section 2 some supporting hyperplane determined by a facet of the polytope $\operatorname{conv}(T - \{v_0\})$ separates v_0 from $\operatorname{conv}(T - \{v_0\})$. This is the desired hyperplane since, by the general position, all the faces of $\operatorname{conv}(T - \{v_0\})$ are simplices.

THEOREM 5.1. *Each m-set in general position in E^d, with $m \geqq d + k + 2$ and $-1 \leqq k \leqq d$, is* $(2, k)$-divisible.

Proof. The theorem is trivial for $d = 1$ and for $k = -1$. For $k = 0$, it reduces to Radon's theorem. It is clearly sufficient to prove the theorem for $m = d + k + 2$. We proceed by induction on both d and k, assuming the statement for $(d + k + 1)$-sets considered either as $[(d-1) + k + 2]$-sets in E^{d-1} or as $[d + (k-1) + 2]$-sets in E^d.

Let T be a $(d + k + 2)$-set in general position in E^d. Pick a point v_0 in $T - \operatorname{conv}(T - \{v_0\})$ and, by the previous lemma, a hyperplane separating v_0 from

$\operatorname{conv}(T - \{v_0\})$ and containing a d-subset of T. Consider the projection

$$\phi:(T - \{v_0\}) \to \pi,$$

defined by $\phi(v_\mu) = v_0v_\mu \cap \pi$. We choose subscripts so that $\phi(v_\mu) = v_\mu$ if and only if $\mu = 1, 2, \cdots, d$; that is, so that v_μ lies in π for these subscripts and only these. The remaining points of T are $v_{d+1}, v_{d+2}, \cdots, v_{d+k+1}$. From the general position it follows that $\phi(v_i) \neq \phi(v_j)$ for $i \neq j$ and $\phi(v_i) \neq v_j$ for $i \neq j$. Hence

$$S = \phi(T - \{v_0\}) = \{v_1, v_2, \cdots, v_d,\ \phi(v_{d+1}), \cdots, \phi(v_{d+k+1})\}$$

is a $(d + k + 1)$-set. We consider two cases:

Case 1. If $k < d$, consider S as a $[(d-1) + k + 2]$-set in general position in the $(d-1)$-space π. By the induction hypothesis there is a partition $\{S_1 S_2\}$ of S with

$$\dim[\operatorname{conv} S_1 \cap \operatorname{conv} S_2] \geqq k.$$

Therefore $\dim \operatorname{conv} S_1 \geqq k$ and S_1 is at least a $(k + 1)$-set. Similarly S_2 is at least a $(k + 1)$-set. A few special cases are simple and instructive as preliminaries.

If $\phi^{-1}(S_1) \cap \pi = \emptyset$, then $S_1 = \{\phi(v_{d+1}), \cdots, \phi(v_{d+k+1})\}$. Observing that

$$\operatorname{conv}[\phi^{-1}(S_1) \cup \{v_0\}] \supset \operatorname{conv} S_1$$

we choose $T_1 = \phi^{-1}(S_1) \cup \{v_0\}$ and $T_2 = S_2$. Then $\{T_1, T_2\}$ is a $(2, k)$-partition of T.

On the other hand, if $\phi^{-1}(S_1) \cap \pi \neq \emptyset$, we must consider several possibilities. Since it is at least a $(k + 1)$-set, S_1 can miss $\{v_1, v_2, \cdots, v_d\}$ only if $S_1 = \{\phi(v_{d+1}), \cdots, \phi(v_{d+k+1})\}$. In this instance we may proceed exactly as in the previous paragraph.

Now consider the general situation and assume that each of S_1 and S_2 contains a v_μ with $1 \leqq \mu \leqq d$. If S_1 contains every $\phi(v_\mu)$, choosing $\{T_1, T_2\}$ as in the previous two situations gives the desired $(2, k)$-partition.

Remaining is the possibility that each of S_1 and S_2 contains both v_μ's with $\mu = 1, 2, \cdots, d$ and $\phi(v_\mu)$'s with $\mu = d + 1, \cdots, d + k + 1$. To deal with this situation we consider two partitions:

$$T_1 = \phi^{-1}(S_1) \cup \{v_0\};\ T_2 = \phi^{-1}(S_2)$$

and

$$T_1' = \phi^{-1}(S_1);\ T_2' = \phi^{-1}(S_2) \cup \{v_0\}.$$

We shall show that one of these must be the desired partition.

First extend the map ϕ in the obvious way to $\operatorname{conv}(T - \{v_0\})$. (To avoid symbol escalation we shall use ϕ to denote both the original map and the extension.) Let $\sigma = \operatorname{conv}\{x_1, x_2, \cdots, x_{k+1}\}$ be a k-simplex with vertices $x_i \in \operatorname{conv} S_1 \cap \operatorname{conv} S_2$. The segment $[\phi^{-1}(x_i)v_0]$ meets both $\operatorname{conv} \phi^{-1}(S_1)$ and $\operatorname{conv} \phi^{-1}(S_2)$. Let x_{i1} and x_{i2}

respectively be the points of intersection with those two sets which are closest to v_0, and let

$$\sigma_1 = \operatorname{conv}\{x_{11}, x_{21}, \cdots, x_{k+1,1}\}$$

$$\sigma_2 = \operatorname{conv}\{x_{12}, x_{22}, \cdots, x_{k+1,2}\}.$$

Since the points x_i were chosen on distinct lines through v_0, σ_1 and σ_2 are k-simplices which map into σ under ϕ.

The special case where $\sigma_1 \cap \sigma_2 = \varnothing$ is instructive. In this situation each x_{i1} lies between x_i and x_{i2} (or vice versa) on the segment $[\phi^{-1}(x_i)v_0]$ (see figure 3). But then

$$\operatorname{conv} T_2' \supset \operatorname{conv}(\sigma_2 \cup \{v_0\}) \supset \sigma_1$$

and, since $\operatorname{conv} T_1' \supset \sigma_1$, we have

$$\dim(\operatorname{conv} T_1' \cap \operatorname{conv} T_2') \geqq \dim \sigma_1 = k.$$

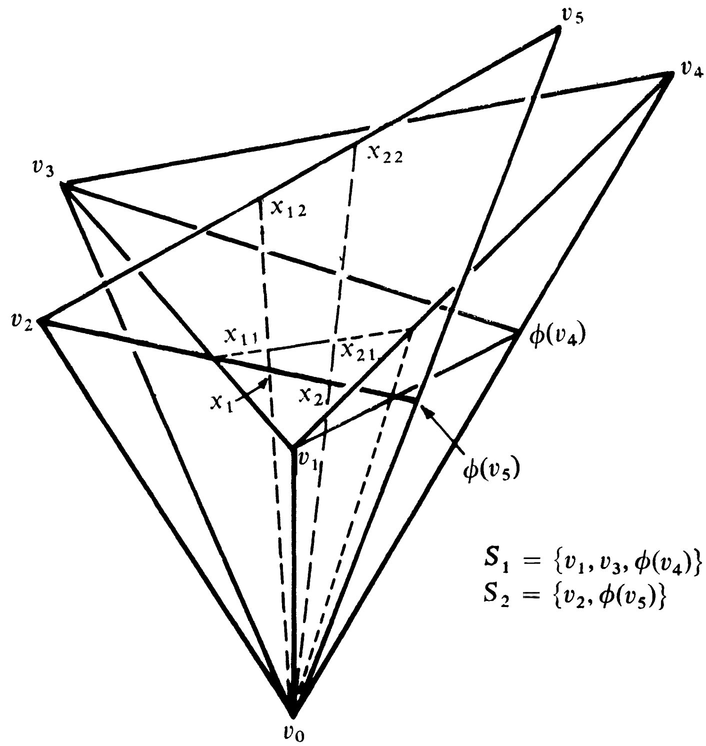


FIG. 3

In general, we break σ_1 into a "lower" set σ_{1l} and an "upper" set σ_{1u} defined by

$y \in \sigma_{1l}$ if the ray $[v_0 y$ meets σ_1 before it meets σ_2.
$y \in \sigma_{1u}$ if $[v_0 y$ meets σ_1 after it meets σ_2 or on σ_2, (see figure 4).

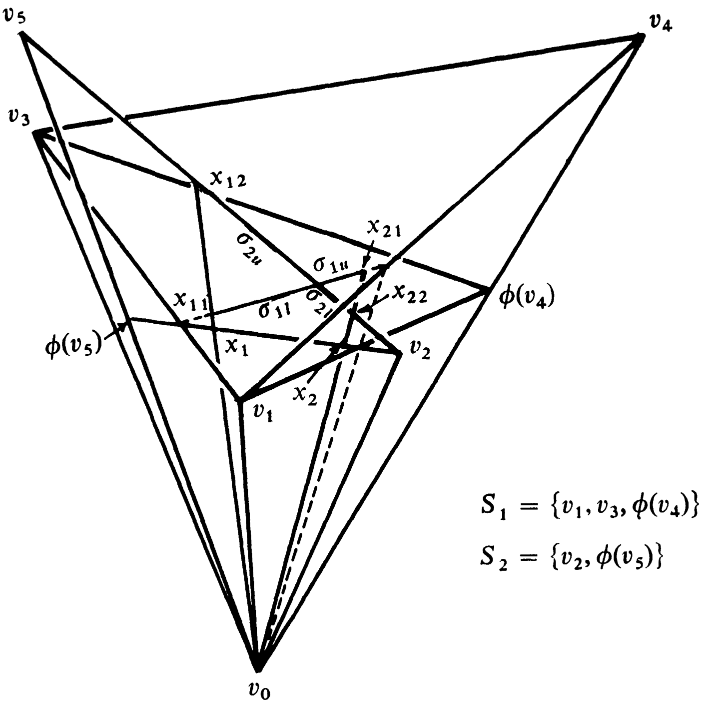


FIG. 4

The set σ_{1l} is open in σ_1 and therefore k-dimensional. Assuming σ_{1l} is not empty (if it is, we can return to the previous case or we can form σ_{2l} and σ_{2u} analogously and can be assured that σ_{2l} is not empty), we have

$$\operatorname{conv} T_2' \supset \operatorname{conv}(\sigma_2 \cup \{v_0\}) \supset \operatorname{conv}(\sigma_{2u} \cup \{v_0\}) \supset \sigma_{1l},$$

and, since $\operatorname{conv} T_1' \supset \sigma_1 \supset \sigma_{1l}$,

$$\dim(\operatorname{conv} T_1' \cap \operatorname{conv} T_2') \geqq \dim \sigma_{1l} = k.$$

Case 2. If $k = d$, so that T is a $(2d+2)$-set, we divide $T - \{v_0\}$ into two disjoint sets

$$T_1^* = \{v_1, v_2, \cdots, v_d\} \quad \text{and} \quad T_2^* = \{v_{d+1}, v_{d+2}, \cdots, v_{2d+2}\}.$$

The set $T - \{v_0\}$ is a $[d + (d-1) + 2]$-set which, by the induction hypothesis, has a $(2, d-1)$-partition $\{T_1, T_2\}$. Letting $\sigma_1 = \operatorname{conv} T_1$ and $\sigma_2 = \operatorname{conv} T_2$ and observing that if $\dim(\sigma_1 \cap \sigma_2) = d$ we are through, assume that $\dim(\sigma_1 \cap \sigma_2) = d - 1$.

One of T_1 and T_2 must be at least a $(d+1)$-set; assume it is T_2. Since T is in general position, $\dim \sigma_2 = d$ and T_1 is at most a d-set. Since $\dim \sigma_1 \geqq d - 1$, we conclude that T_1 is exactly a d-set and T_2 exactly a $(d+1)$-set.

If $T_1 = T_1^*$, then $T_2 = T_2^*$ and $\operatorname{conv} T_1 \cap \operatorname{conv} T_2 = \emptyset$. Therefore $T_1 \cap T_2^* \neq \emptyset$ and $T_2 \cap T_1^* \neq \emptyset$. Moreover, since T_2 is a $(d+1)$-set, $T_2 \cap T_2^* \neq \emptyset$.

If $\sigma_1 \cap \sigma_2$ lies in $\operatorname{Bd}\sigma_2$, it is contained in a $(d-1)$-simplex on $\operatorname{Bd}\sigma_2$. The hyperplane determined by this simplex must contain a d-subset of T_2 and the entire d-set T_1. But then we have $2d$ points of T in a hyperplane, contradicting the general position.

Hence σ_1 meets σ_2 on its interior. We choose a partition of T:

$$T_1' = T_1 \cup \{v_0\}; \; T_2' = T_2.$$

Since v_0 is not on the hyperplane $\operatorname{aff}\sigma_1$, we have

$$\dim(\operatorname{conv} T_1' \cap \operatorname{conv} T_2') = \dim[\operatorname{conv}(\sigma_1 \cup \{v_0\}) \cap \sigma_2] \geqq d.$$

This completes the proof.

A Radon partition $\{T_1, T_2\}$ is of **type** $\{r,s\}$ if T_1 is an r-set and T_2 an s-set. A natural question is, "Given a k-set T, what types of Radon partitions can occur?" Given positive integers r and s such that $r + s = d + 2$, one can construct without difficulty a $(d+2)$-set T in general position whose Radon partition (unique by Corollary 4.1) is of type $\{r,s\}$.

For some sets, the existence of certain types of partitions can easily be ruled out. A polytope P is **k-neighbourly** if every k-subset R of $\operatorname{vert} P$ determines a proper face $F = \operatorname{conv} R$ of P. Neighbourly polytopes are unfamiliar to many since no examples other than simplices exist in dimensions two and three. The cyclic polytopes [**5**, sec. 4.7] however, provide examples of $[d/2]$-neighbourly d-polytopes with arbitrarily many vertices. Moreover, this is the best we can hope for since a d-polytope which is k-neighbourly for any $k > [d/2]$ must in fact be a simplex [**5**, sec. 7.1]. If $P = \operatorname{conv} T$ is k-neighbourly, then it is r-neighbourly for $r \leqq k$. Hence if $T = \operatorname{vert} P$ and $T_1 \subset T$ is an r-set, $\operatorname{conv} T_1$ is a proper face of P. But then there must be a supporting hyperplane π to P such that $\pi \cap T = T_1$, so that T_1 cannot be a member of any Radon partition. Therefore T admits no Radon partition of type $\{r,s\}$ for $r \leqq k$ or $s \leqq k$. This result is included in the following theorem of Shephard [**18**]:

THEOREM 5.2. *Let T be an m-set in E^d with $m \geqq d + 3$. Then*:

(i) *If T is not the vertex set of a polytope, it admits a Radon partition of type $\{r, m-r\}$ for $1 \leqq r \leqq m-1$.*

(ii) *If T is the vertex set of a k-neighbourly polytope, it admits no Radon partition of type $\{r, m-r\}$ for $r \leqq k$ or $r \geqq m-k$.*

(iii) *If T is the vertex set of a polytope which is k-neighbourly but not $(k+1)$-neighbourly, it admits a Radon partition of type $\{r, m-r\}$ for $k+1 \leqq r \leqq m-k-1$.*

6. The theorems of Helly and Caratheodory and related results. Inductive methods employing the separation lemmas and projections onto hyperplanes can be employed to prove Helly's theorem without difficulty. To avoid repetition and llustrate an elegant use of Radon's theorem we choose a more familiar proof.

THEOREM 6.1. (Helly) *If $C_1, C_2, \cdots, C_k$ are convex sets in E^d, any $d+1$ of which have non-empty intersection, then* $\bigcap_{\mu=1}^{k} C_\mu \neq \varnothing$.

Proof. Consider the first interesting case, where $k = d+2$. By hypothesis, we can find points

$$v_j \in \bigcap_{\mu \neq j} C_\mu \quad \text{for } j = 1, 2, \cdots, d+2.$$

By Radon's theorem the set $T = \{v_1, v_2, \cdots, v_{d+2}\}$ admits a partition $\{T_1, T_2\}$ with $\operatorname{conv} T_1 \cap \operatorname{conv} T_2 \neq \varnothing$. If $v_j \in T_1$, then $\operatorname{conv} T_2 \subset C_j$. If $v_j \in T_2$, then $\operatorname{conv} T_1 \subset C_j$. Therefore

$$\bigcap_{\mu=1}^{d+2} C_\mu \supset \operatorname{conv} T_1 \cap \operatorname{conv} T_2 \neq \varnothing.$$

The proof is completed by a straightforward induction on k, considering the sets $C'_j = C_j \cap C_{k+1}$ for $j = 1, 2, \cdots, k$.

Helly's theorem does not hold for infinite collections without further hypotheses; compactness of the sets being the most obvious one.

Probably the best extension of Helly's theorem is the following result of Horn [7] and Klee [9]. We offer a proof of one part which succumbs easily to the geometric approach via projections on hyperplanes.

THEOREM 6.2. *If $F = \{C_i\}$ is a family of compact convex sets in E^d and $1 \leqq k \leqq d+1$, then the following are equivalent:*

(i) *Every k of the sets C_i have non-empty intersection.*

(ii) *Each $(d-k)$-flat in E^d lies in a $(d-k+1)$-flat in E^d which meets every member of F.*

(iii) *Each $(d-k+1)$-flat has a translate which meets every member of F.*

Proof. (i) implies (iii). Let π be a $(d-k+1)$-flat and π' an orthogonal $(k-1)$-flat. Project E^d onto π' parallel to π. Convexity, compactness, and intersections are preserved, so that the images satisfy the hypothesis of Helly's theorem. The $(d-k+1)$-flat, parallel to π and containing a point common to all the images, meets every member of the family F and is the desired flat.

An argument from the separation lemmas may be used to establish Caratheodory's theorem. We choose here a slightly different, but still purely geometric, approach. Our proof of the succeeding theorem, the important generalization of Caratheodory's result due to Steinitz [**19**], will again make strong use of projections on hyperplanes.

THEOREM 6.2. (Caratheodory) *If $T \subset E^d$ and $x \in \operatorname{conv} T$, then there is an (at most) $(d+1)$-set $S \subset T$ with $x \in \operatorname{conv} S$.*

Proof. We first show that $\operatorname{conv} T$ is in fact the union of the sets $\operatorname{conv} S$, where S is a *finite* subset of T. If S_1 and S_2 are finite subsets of T with $x_1 \in \operatorname{conv} S_1$ and $x_2 \in \operatorname{conv} S_2$, then $S_1 \cup S_2$ is a finite subset of T and $[x_1 x_2] \subset \operatorname{conv}(S_1 \cup S_2)$.

Hence the set

$$T^* = \cup \{\operatorname{conv} S \mid S \text{ finite}, S \subset T\}$$

is convex and contains T. Since $\operatorname{conv} S \subset \operatorname{conv} T$, we have $T^* = \operatorname{conv} T$.

The proof now proceeds by induction on the dimension d. For $d = 1$ the statement is trivial. Assume it proved for $d - 1$ and pick a finite set $S_1 \subset T$ with $x \in \operatorname{conv} S_1$. If $x \in \operatorname{Bd} \operatorname{conv} S_1$, there is a hyperplane π supporting $\operatorname{conv} S_1$ at x and $(\operatorname{conv} S_1) \cap \pi = \operatorname{conv}(S_1 \cap \pi)$. Applying the induction hypothesis in π, there is an (at most) d-set $S_2 \subset S_1 \cap \pi \subset T$ with $x \in \operatorname{conv} S_2$.

Otherwise we can pick a point $y \in S_1$ and let $z = [yx \cap \operatorname{Bd} \operatorname{conv} S_1]$. There must be such a point because S_1 is finite. By the previous argument there is an (at most) d-set $S_2 \subset T$ with $z \in \operatorname{conv} S_2$. The set $S = S_2 \cup \{y\}$ is at most a $(d + 1)$-set and $x \in \operatorname{conv} S$.

THEOREM 6.3 (Steinitz). *If $T \subset E^d$ and $x \in \operatorname{int} \operatorname{conv} T$, then there is an (at most) $2d$-set $S \subset T$ with $x \in \operatorname{int} \operatorname{conv} S$.*

Proof. Again the statement is trivial for $d = 1$ and we proceed by induction on d. Let π be a hyperplane on x.

Since $x \in \operatorname{int} \operatorname{conv} T$, there are points v_0 and v_1 of T on opposite sides of π. Consider the mapping $\phi: T \to \pi$ defined by

$$\phi(y) = [v_0 y] \cap \pi \quad \text{or} \quad [v_1 y] \cap \pi.$$

This mapping is well defined because the only points for which both $[v_0 y]$ and $[v_1 y]$ meet π are v_0, v_1, and the points of $T \cap \pi$. Clearly $x \in \operatorname{relint} \operatorname{conv} \phi(T)$, the interior of $\operatorname{conv} \phi(T)$ relative to the hyperplane π. By the induction hypothesis there is a $2(d - 1)$-set $S_1 \subset \phi(T)$ with $x \in \operatorname{relint} \operatorname{conv} S_1$. For each $y \in S_1$ we pick a single representative from the set $\phi^{-1}(y)$, and denote the collection of all such points by S_2. The set $S = S_2 \cup \{v_0\} \cup \{v_1\}$ is at most a $2(d - 1) + 2 = 2d$-subset of T. Since $S_1 \subset \operatorname{conv} S$, we have $x \in \operatorname{int} \operatorname{conv} S$. This completes the proof.

The methods used here could probably be exploited to supply more geometric proofs of related results due to Robinson [**17**], Bonnice-Klee [**2**], and Reay [**15**]. It is also probable, however, that the increasing complexity of the separations and projections involved would tend to obscure the underlying situation we have tried to expose.

References

1. B. J. Birch, On $3N$ points in a plane, Proc. Cambridge Philos. Soc., 55 (1959) 289–293.

2. W. Bonnice and V. L. Klee, The generation of convex hulls, Math. Ann., 152 (1963) 1–29.

3. G. D. Chakerian, Intersection and covering properties of convex sets, this MONTHLY, 76 (1969) 753–766.

4. L. Danzer, B. Grünbaum, and V. Klee, Helly's Theorem and its relatives, Amer. Math. Soc., Proc. Sympos. Pure Math., 7(1963) 101–180.

5. B. Grünbaum, Convex Polytopes, Wiley, New York, 1967.

6. W. Gustin, On the interior of the convex hull of a euclidean set, Bull. Amer. Math. Soc., 53 (1947) 299–301.

7. A. Horn, Some generalizations of Helly's theorem on convex sets, Bull. Amer. Math. Soc., 55 (1949) 923–929.

8. J. Kenelly and W. Hare, Characterizations of Radon partitions, Pacific J. Math, (to appear).

9. V. Klee, On certain intersection properties of convex sets, Canad. J. Math., 3 (1951) 272–275.

10. L. Kosmak, A remark on Helly's theorem, Spisy Prirod Fak, Univ. Brno, 1963, 223–225.

11. I. V. Proskuryakov, A property of n-dimensional affine space connected with Helly's Theorem, Usp. Math. Nauk., 14 (1959) 219–222.

12. R. Rado, Theorems on the intersection of convex sets of points, J. London Math. Soc., 27 (1952) 320–328.

13. J. Radon, Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten, Math. Ann., 83 (1921) 113–115.

14. J. R. Reay, An extension of Radon's theorem, Illinois J. Math., 12 (1968) 184–189.

15. ———, Generalizations of a theorem of Caratheodory, Memoirs Amer. Math. Soc., No. 54, Providence, 1965.

16. H. Rademacher and I. J. Schoenberg, Helly's theorem on convex domains and Tchebycheff's approximation problem, Canad. J. Math., 2 (1950) 245–256.

17. C. V. Robinson, Spherical theorems of Helly type and congruence indices of spherical caps, Amer. J. Math., 64 (1942) 260–272.

18. G. C. Shephard, Neighbourliness and Radon's theorem, Mathematika, 16 (1969) 273–275.

19. E. Steinitz, Bedingt konvergente Reihen und konvexe Systeme I–II–III, J. Reine Angew. Math., 143 (1913) 128–175; 144 (1914) 1–40; 146 (1916) 1–52.

20. H. Tverberg, A generalization of Radon's theorem, J. London Math. Soc., 41 (1966) 123–128.

INTEGRATION IN FINITE TERMS

MAXWELL ROSENLICHT, University of California, Berkeley

1. The question arises in elementary calculus: Can the indefinite integral of an explicitly given function of one variable always be expressed "explicitly" (or "in closed form", or "in finite terms")? Liouville gave the answer one would expect, "No", and he proved in particular that such is not the case with $\int e^{x^2}dx$. Since we have all fallen into the habit of quoting this result and giving neither proof nor reference, it may be worthwhile to actually state it as precisely as possible and give a proof that is as elementary as the subject matter might suggest.

We must define our terms carefully. To begin with, we are not interested in arbitrary functions, but in **elementary functions,** which are functions of one variable

Maxwell Rosenlicht did his Ph.D. work at Harvard, under Oscar Zariski. He was a National Research Fellow at Chicago and Princeton, and held a position at Northwestern Univ. before his present position at the Univ. of California, Berkeley. He has spent visits at the Univ. of Rome, the IHES-Paris, the Univ. of Mexico, Harvard Univ., and Northwestern Univ., and he has had Fulbright and Guggenheim Fellowships. In 1960, he won the American Mathematical Society's Cole Prize in Algebra (with S. Lang). His main research is in algebraic geometry, algebraic groups, and differential algebra, and he is the author of *Introduction to Analysis* (Scott, Foresman, 1968). *Editor.*

built up by using that variable and constants, together with repeated algebraic operations and the taking of exponentials and logarithms. Since we lose no generality by doing so, we shall take all exponentials and logarithms to the base e. We allow ourselves the convenience of the use of complex numbers, for with these the various trigonometric and inverse trigonometric functions turn out to be elementary, as seems reasonable. Thus the integral of a rational function of one real variable is elementary, since it is a linear combination of logarithms, inverse tangents, and rational functions. But we are still deficient in precision, because of the multivaluedness of algebraic functions and logarithms. The functions we work with must be specific objects, each susceptible of an unambiguous sense. We choose to avoid the difficulties associated with multivaluedness by the simplest method, that of restricting ourselves, in any given discussion, to functions on some specific region (that is, nonempty connected open subset) of the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$, and furthermore considering only meromorphic functions on the region in question, a meromorphic function on a region being a function whose values are complex numbers or the symbol ∞, with the property that sufficiently near any point z_0 of the region the function is given by a convergent Laurent series in $z - z_0$, that is, a convergent power series in $z - z_0$, with the possible addition of a finite number of negative powers. Thus the rational functions of one variable, which form the field $\mathbb{C}(z)$ got by adjoining the identity function z to the field of constant functions $\mathbb{C}$, are all meromorphic on all of $\mathbb{R}$ or $\mathbb{C}$. The exponential of a function f meromorphic on a certain region of $\mathbb{R}$ or $\mathbb{C}$ is a function meromorphic on the subregion obtained by deleting those points where the value of f is ∞ (and then taking a connected component, if we are working in $\mathbb{R}$), while $\log f$ can be taken to be meromorphic on any simply connected subregion where f takes on neither of the values 0 or ∞, by arbitrarily choosing one of its many values at any particular point of the subregion. Furthermore, the implicit function theorem shows that if we are given a polynomial equation with coefficients which are functions meromorphic on a certain region, the leading coefficient not being zero, then there exists a meromorphic solution on a suitable subregion. Thus any complicated expression for an elementary function, compounded of algebraic operations, exponentials and logarithms, has a realization as a meromorphic function on some region. Now the totality of all meromorphic functions on a given region form a field under the usual operations of functional addition and multiplication, and the restriction of all these functions to any given subregion gives an embedding of fields. The derivative of a function meromorphic on a given region is again meromorphic, as is an indefinite integral, if one exists, of the function. Note that the rational functions on a region, that is the restriction of $\mathbb{C}(z)$ to this region, are a field of meromorphic functions on the region that are closed under differentiation, and that if we have any field of meromorphic functions on a region that is closed under differentiation and get a larger field of meromorphic functions on the region by adjoining the exponential or a logarithm of a function in our field, or a solution

of a polynomial equation with coefficients in the field, we again get a field of meromorphic functions on the region that is closed under differentiation. Thus the proper objects of study are seen to be fields of meromorphic functions on given regions in $\mathbb{R}$ or $\mathbb{C}$ which are closed under differentiation. If a function in such a field has an indefinite integral that is expressible "in finite terms," then by restricting all functions, if necessary, to a suitable subregion, we see that we have a tower of such fields of meromorphic functions, each larger field being obtained by adjunction of an exponential, or a logarithm, or the solution of an algebraic equation, the tower starting with the original field and culminating in a field containing the indefinite integral. Thus the original loosely worded analytic problem, when formulated as a precise analytic problem, becomes algebraic.

2. Define a **differential field** to be a field F, together with a **derivation** on F, that is, a map of F into itself, usually denoted $a \mapsto a'$, such that $(a + b)' = a' + b'$ and $(ab)' = a'b + ab'$ for all $a, b \in F$. Immediate consequences are that $(a/b)' = (ab' - a'b)/b^2$ if a, $b \in F$, $b \neq 0$, and $(a^n)' = na^{n-1}a'$ for all integers n. Furthermore, $1' = (1^2)' = 2 \cdot 1 \cdot 1'$, so $1' = 0$. Therefore the **constants** of F, that is, all $c \in F$ such that $c' = 0$, are a subfield of F.

If a, b are elements of the differential field F, a being nonzero, let us agree to call a **an exponential of** b, or b **a logarithm of** a, if $b' = a'/a$; this terminology is not unreasonable for our present purposes since the only properties of exponentials and logarithms in which we are interested are their differential properties. We immediately get the "logarithmic derivative identity,"

$$\frac{(a_1^{\nu_1} \cdots a_n^{\nu_n})'}{a_1^{\nu_1} \cdots a_n^{\nu_n}} = \nu_1 \frac{a_1'}{a_1} + \cdots + \nu_n \frac{a_n'}{a_n},$$

for $a_1, \cdots, a_n$ nonzero elements of F and $\nu_1, \cdots, \nu_n$ integers.

3. There is a standard result on algebraic extensions of differential fields which we shall need later. For completeness we prove it here. The result is that if F is a differential field of characteristic zero and K an algebraic extension field of F, then the derivation on F can be extended to a derivation on K, and this extension is unique. (Thus K has a unique differential field structure extending that of F. We remark that the restriction to characteristic zero is not essential; it suffices to assume that K is separable over F, and the following proof will hold in this more general case.) For the reader who is interested only in the classical function-theoretic case, where the fields in question are fields of meromorphic functions on a region of $\mathbb{R}$ or $\mathbb{C}$, the proof is immediate, the existence proof being a direct consequence of the implicit function theorem, uniqueness following from the ordinary method of computing derivatives of functions given implicitly. To prove the result generally, let X be an indeterminate and define the maps D_0, D_1 of the polynomial ring $F[X]$ into itself by

$$D_0\left(\sum_{i=0}^{n} a_iX^i\right) = \sum_{i=0}^{n} a_i'X^i, \quad D_1\left(\sum_{i=0}^{n} a_iX^i\right) = \sum_{i=0}^{n} ia_iX^{i-1}$$

for $a_0, a_1, \cdots, a_n \in F$. If K has a differential field structure extending that of F, then for any $x \in K$ and any $A(X) \in F[X]$ we have

$$(A(x))' = (D_0A)(x) + (D_1A)(x) \cdot x'.$$

If we replace $A(X)$ by the minimal polynomial $f(X)$ of x over F, (that is, the monic irreducible polynomial of which x is a root, indeed a simple root, so that $(D_1f)(x) \neq 0$), we get $x' = -(D_0f)(x)/(D_1f)(x)$. Thus the differential field structure on K that extends that on F is unique, if it exists. We now show that such a structure on K exists. Using the usual field-theoretic arguments, we may assume that K is a finite extension of F, so that we can write $K = F(x)$, for a certain $x \in K$. For some $g(X) \in F[X]$, to be determined later, let the map $D: F[X] \to F[X]$ be defined by

$$DA = D_0A + g(X)D_1A,$$

for any $A \in F[X]$. It follows immediately that $D(A + B) = DA + DB$ and $D(AB) = (DA)B + A(DB)$ for all $A, B \in F[X]$, since the analogous identities hold for both D_0 and D_1. Note that $Da = a'$ for all $a \in F$. Now look at the natural surjective ring homomorphism $F[X] \to F[x]$, which is the identity on F and sends X into x. Since $F[x] = F(x) = K$, the map D on $F[X]$ will induce a derivation on K extending that on F if it so happens that D maps the kernel of our ring homomorphism into itself. But the kernel of the homomorphism is the ideal $F[X]f(X)$, where $f(X)$ is the minimal polynomial of x over F. Hence we shall have proved our result once we have shown that D maps $F[X]f(X)$ into itself. The condition for this is simply that D map $f(X)$ into a multiple of itself, that is that Df be any element of $F[X]$ of which x is a root, or that $(Df)(x) = 0$. But this last condition reduces to $(D_0f)(x) + g(x)(D_1f)(x) = 0$. Since $(D_1f)(x) \neq 0$ and $F(x) = F[x]$, a polynomial $g(X) \in F[X]$ can actually be found such that $(Df)(x) = 0$, and this completes the proof of our statement.

4. By a **differential extension field** of a differential field F we mean, of course, a differential field which is an extension field of F whose derivation extends the derivation on F. The following result will be the principal tool for proving the theorem of the next section, and will be used for the verification of our subsequent examples.

LEMMA. *Let F be a differential field, $F(t)$ a differential extension field of F having the same subfield of constants, with t transcendental over F, and with either $t' \in F$ or $t'/t \in F$. If $t' \in F$, then for any polynomial $f(t) \in F[t]$ of positive degree, $(f(t))'$ is a polynomial in $F[t]$ of the same degree as $f(t)$, or degree one less, accrdding as the highest coefficient of $f(t)$ is not, or is, a constant. If $t'/t \in F$, then for any nonzero $a \in F$ and any nonzero integer n we have $(at^n)' = ht^n$, for some nonzero*

$h \in F$, and furthermore, for any polynomial $f(t) \in F[t]$ of positive degree, $(f(t))'$ is a polynomial in $F[t]$ of the same degree, and is a multiple of $f(t)$ only if $f(t)$ is a monomial.

We first consider the case $t' = b \in F$. Let the degree of $f(t)$ be $n > 0$, so that $f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_0$, with $a_0, \cdots, a_n \in F$, $a_n \neq 0$. Then

$$(f(t))' = a_n' t^n + (na_n b + a_{n-1}')t^{n-1} + \cdots .$$

This is clearly a polynomial in $F[t]$, of degree n if a_n is not constant. If a_n is constant and $na_n b + a_{n-1}' = 0$, then $(na_n t + a_{n-1})' = na_n b + a_{n-1}' = 0$, so that $na_n t + a_{n-1}$ is a constant, therefore an element of F, so that $t \in F$, contrary to the assumption that t is transcendental over F. Thus if a_n is constant, $(f(t))'$ has degree $n - 1$.

Now suppose that we are in the case $t'/t = b \in F$. Let $a \in F$, $a \neq 0$, and let n be a nonzero integer. Then

$$(at^n)' = a't^n + nat^{n-1}t' = (a' + nab)t^n.$$

If $a' + nab = 0$, then $(at^n)' = 0$, so that at^n is constant, therefore an element of F, contradicting the transcendence of t over F. Therefore $a' + nab \neq 0$. Finally, let $f(t) \in F[t]$ have positive degree. Clearly $(f(t))'$ has the same degree. If $(f(t))'$ is a multiple of $f(t)$, it must be by a factor in F. Therefore if $f(t)$ is not a monomial, $a_n t^n$ and $a_m t^m$ being two of its different terms, and $(f(t))'$ is a multiple of $f(t)$, we have

$$\frac{a_n' + na_n b}{a_n} = \frac{a_m' + ma_m b}{a_m},$$

so

$$\frac{a_n'}{a_n} + n\frac{t'}{t} = \frac{a_m'}{a_m} + m\frac{t'}{t},$$

or $(a_n t^n / a_m t^m)' = 0$, so that $a_n t^n / a_m t^m \in F$, again contradicting the transcendence of t over F. This completes the proof.

5. Let F be a differential field. Define an **elementary extension of** F to be a differential extension field of F which is obtained by successive adjunctions of elements that are algebraic, or logarithms, or exponentials, that is, a differential extension field of the form $F(t_1, \cdots, t_N)$, where for each $i = 1, \cdots, N$, the element t_i is either algebraic over the field $F(t_1, \cdots, t_{i-1})$, or the logarithm or exponential of an element of $F(t_1, \cdots, t_{i-1})$. Note that each intermediate field $F(t_1, \cdots, t_{i-1})$ is a differential field and an elementary extension of F.

The following result is the abstract generalization of Ostrowski's 1946 generalization of Liouville's 1835 theorem on the subject. A proof of the analytic case may be found in Ritt's classic exposition [**4**]. Other algebraic proofs, essentially the same as the one given here, may be seen in [**2**] and [**5**].

THEOREM. *Let F be a differential field of characteristic zero and $\alpha \in F$. If the equation $y' = \alpha$ has a solution in some elementary differential extension field of F having the same subfield of constants, then there are constants $c_1, \cdots, c_n \in F$ and elements $u_1, \cdots, u_n, v \in F$ such that*

$$\alpha = \sum_{i=1}^{n} c_i \frac{u_i'}{u_i} + v'.$$

A number of comments are in order before we proceed with the proof. First, in the case of greatest interest, in which our fields are fields of meromorphic functions on some subregion of $\mathbb{R}$ or $\mathbb{C}$, the condition that F and its elementary extension field have the same constants will be automatically satisfied as long as $\mathbb{C} \subset F$, since any constant meromorphic function is a complex number. In the general case however, the condition that F and its elementary extension field have the same constants, or some related condition, is essential. This can be seen from the example $F = \mathbb{R}(x)$, the field of real rational functions of a real variable, with $x' = 1$ as usual, and $\alpha = 1/(x^2 + 1)$. Clearly $\int (1/(x^2 + 1))dx$ is an element of an elementary extension field of $\mathbb{R}(x)$, and our claim is that the assumption that we can write $1/(x^2 + 1)$ in the desired form, with $c_1, \cdots, c_n \in \mathbb{R}$ and $u_1, \cdots, u_n$, $v \in \mathbb{R}(x)$, will lead to a contradiction. For if $x^2 + 1$ occurs v_i times in the expression of u_i as a power product of monic irreducible elements of $\mathbb{R}[X]$, then $u_i'/u_i - 2v_i x/(x^2 + 1)$ is an element of $\mathbb{R}(x)$ without $x^2 + 1$ in its denominator, while $x^2 + 1$, if it occurs in the denominator of v, will occur at least twice in the denominator of v'. Thus $x^2 + 1$ divides the denominator of neither v nor v', implying that $1 - \Sigma 2c_i v_i x$ is divisible by $x^2 + 1$, which is impossible. The final comment is that the theorem has an easy converse: if α can be written as indicated then α has an integral in some elementary extension field of F. This is quite easy to show in the abstract case and is immediate in the classical case where F is a field of meromorphic functions on a subregion of $\mathbb{R}$ or $\mathbb{C}$, as we see by passing to a suitable subregion, where the various $\log u_i$'s can be defined.

Now for the proof of Liouville's theorem. By assumption there is a tower of differential fields

$$F \subset F(t_1) \subset \cdots \subset F(t_1, \cdots, t_N),$$

all with the same subfield of constants, each t_i being algebraic over $F(t_1, \cdots, t_{i-1})$, or the logarithm or exponential of an element of this field, such that there exists an element $y \in F(t_1, \cdots, t_N)$ such that $y' = \alpha$. We shall prove the theorem by induction on N. The case $N = 0$ is trivial, so assume that $N > 0$ and that the theorem holds for $N - 1$. Applying the case $N - 1$ to the fields $F(t_1) \subset F(t_1, \cdots, t_N)$, we deduce that we can write α in the desired form, but with $u_1, \cdots, u_n, v$ in $F(t_1)$. Setting $t_1 = t$, we have t algebraic over F, or the logarithm or exponential of an element of F, and we know that

$$\alpha = \sum_{i=1}^{n} c_i \frac{u_i'}{u_i} + v',$$

with $c_1, \cdots, c_n$ constants of F and $u_1, \cdots, u_n$, $v \in F(t)$, and it remains to find a similar expression for α, possibly with a different n, but with all of $u_1, \cdots, u_n, v$ in F.

First suppose that t is algebraic over F. Then there are polynomials $U_1, \cdots, U_n$, $V \in F[X]$ such that $U_1(t) = u_1, \cdots, U_n(t) = u_n$, $V(t) = v$. Let the distinct conjugates of t over F in some suitable algebraic closure of $F(t)$ be $\tau_1\,(=t)$, $\tau_2, \cdots, \tau_s$. (In case we are dealing with fields of meromorphic functions on a region in $\mathbb{R}$ or $\mathbb{C}$, the functions $\tau_2, \cdots, \tau_s$ can be taken to be meromorphic functions on a suitable subregion, and it suffices to carry the proof through for functions on the subregion.) Now bear in mind the result of Section 3 on algebraic extensions of differential fields. We have

$$\alpha = \sum_{i=1}^{n} c_i \frac{(U_i(\tau_j))'}{U_i(\tau_j)} + (V(\tau_j))'$$

for $j = 1, \cdots, s$, since this is true for $j = 1$. Application of the operation $(1/s)\sum_{j=1}^{s}$ to both sides of the equation yields

$$\alpha = \sum_{i=1}^{n} \frac{c_i}{s} \frac{(U_i(\tau_1)\cdots U_i(\tau_s))'}{U_i(\tau_1)\cdots U_i(\tau_s)} + \left(\frac{V(\tau_1)+\cdots+V(\tau_s)}{s}\right)'.$$

Since each $U_i(\tau_1)\cdots U_i(\tau_s)$ and $V(\tau_1)+\cdots+V(\tau_s)$ are symmetric polynomials in $\tau_1, \cdots, \tau_s$ with coefficients in F, each of these expressions is actually in F. Hence the last equation is an expression for α of the desired form.

In the remaining cases, where t is the logarithm or exponential of an element of F, we may assume that t is transcendental over F. Then we have

$$\alpha = \sum_{i=1}^{n} c_i \frac{(u_i(t))'}{u_i(t)} + (v(t))',$$

with $u_1(t), \cdots, u_n(t)$, $v(t) \in F(t)$. Each $u_i(t)$ can be written as a power product of a nonzero element of F and various monic irreducible elements of $F[t]$. Hence we may, if necessary, use the logarithmic derivative identity to rewrite $\sum c_i(u_i(t))'/u_i(t)$ in a similar form, but with each $u_i(t)$ either in F or a monic irreducible element of $F[t]$. We therefore assume that $u_1(t), \cdots, u_n(t)$ are distinct, each being an element of F or a monic irreducible element of $F[t]$, and that no c_i is zero. Now look at the partial fraction decomposition of $v(t)$, which expresses $v(t)$ as the sum of an element of $F[t]$ plus various terms of the form $g(t)/(f(t))^r$, where $f(t)$ is a monic irreducible element of $F[t]$, r a positive integer, and $g(t)$ is a nonzero element of $F[t]$ of degree less than that of $f(t)$. Clearly $u_1(t), \cdots, u_n(t), v(t)$ must be of very special form for the right hand side of the last equation to add up to α, which doesn't involve t. To investigate this special form in detail, it now becomes convenient to separate cases. In each case the lemma provides the basic arguments.

First, suppose that t is the logarithm of an element of F, so that $t' = a'/a$, for some $a \in F$. Let $f(t)$ be a monic irreducible element of $F[t]$. Then $(f(t))'$ is also in $F[t]$, and it has degree less than that of $f(t)$, so that $f(t)$ does not divide $(f(t))'$.

Thus if $u_i(t) = f(t)$, then the fraction $(u_i(t))'/u_i(t)$ is already in lowest terms, with denominator $f(t)$. If $g(t)/(f(t))^r$ occurs in the partial fraction expression for $v(t)$, with $g(t) \in F[t]$ of degree less than that of $f(t)$ and $r > 0$ and maximal for given $f(t)$, then $(v(t))'$ will consist of various terms having $f(t)$ in the denominator at most r times plus $(g(t)(1/(f(t))^r)' = -rg(t)(f(t))'/(f(t))^{r+1}$. Since $f(t)$ does not divide $g(t)(f(t))'$, we see that a term with denominator $(f(t))^{r+1}$ actually appears in $(v(t))'$. Thus if $f(t)$ appears as a denominator in the partial fraction expansion of $v(t)$, it will appear in α, which is impossible. Therefore, $f(t)$ does not appear in the denominator of $v(t)$. Therefore $f(t)$ cannot be one of the $u_i(t)$'s either. Since this is true for each monic irreducible $f(t)$, we have each $u_i(t) \in F$ and $v(t) \in F[t]$. Since $(v(t))' \in F$, the lemma implies that $v(t) = ct + d$, with c constant and $d \in F$. Thus

$$\alpha = \sum_{i=1}^{n} c_i \frac{u_i'}{u_i} + c\frac{a'}{a} + d'$$

is an expression for α of the desired form.

Finally, consider the case where t is the exponential of an element of F, say $t'/t = b'$, with $b \in F$. The lemma implies that if $f(t)$ is a monic irreducible element of $F[t]$ other than t itself, then $(f(t))' \in F[t]$ and $f(t)$ does not divide $(f(t))'$. Precisely the same reasoning as above shows that $f(t)$ cannot occur in the denominator of $v(t)$, nor can any $u_i(t)$ equal $f(t)$. Thus $v(t)$ can be written as $v(t) = \sum_j a_j t^j$, where each $a_j \in F$ and j ranges over a finite set of integers, positive, negative, or zero, and each of the quantities $u_1(t), \cdots, u_n(t)$ is in F, with the possible exception that one of these may be t itself. Since each $(u_i(t))'/u_i(t)$ is in F, we have $(v(t))' \in F$, so the lemma implies that $v(t) \in F$. If each $u_i(t)$ is in F, we already have α in the desired form, and are done. If not, only one $u_i(t)$, say $u_1(t)$, is not in F. Then $u_1(t) = t$ and $u_2(t), \cdots, u_n(t) \in F$, so we can write

$$\alpha = c_1 \frac{t'}{t} + \sum_{i=2}^{n} c_i \frac{u_i'}{u_i} + v' = \sum_{i=2}^{n} c_i \frac{u_i'}{u_i} + (c_1 b + v)',$$

with $u_2, \cdots, u_n$, $c_1 b + v$ all in F. This completes the proof of the theorem.

6. An elementary function is a meromorphic function on some region in $\mathbb{R}$ or $\mathbb{C}$ that is contained in an elementary extension field of the field of rational functions $\mathbb{C}(z)$. We now give some examples of elementary functions with nonelementary indefinite integrals.

As a preliminary comment we note that if $g(z)$ is a non-constant rational function of the complex variable z then e^g is not algebraic over $\mathbb{C}(z)$. This can easily be shown analytically by noting that since $g(z)$ must have at least one pole on the Riemann sphere, e^g will have at least one essential singularity, unlike any algebraic function. Or it can be shown algebraically by looking at the irreducible equation over $\mathbb{C}(z)$ that e^g would otherwise satisfy, say

$$e^{ng} + a_1 e^{(n-1)g} + \cdots + a_n = 0,$$

where $a_1, \cdots, a_n \in \mathbb{C}(z)$, then differentiating this to get

$$ng'e^{ng} + (a_1' + (n-1)a_1 g')e^{(n-1)g} + \cdots + a_n' = 0,$$

which must be proportional to the first equation, so that $ng' = a_n'/a_n$, then noting that a_n'/a_n is either zero or a sum of fractions with constant numerators and linear denominators, whereas ng' can have no linear denominator, so that $g' = 0$, contradicting the assumption that g is nonconstant.

We now want to derive a criterion, due to Liouville, that $\int f(z)e^{g(z)}dz$ be elementary, where $f(z), g(z)$ are given rational functions of z, $f(z)$ being nonzero, and $g(z)$, as above, non-constant. Writing $e^g = t$, we have $t'/t = g'$. Working in the differential field $\mathbb{C}(z,t)$, a pure transcendental extension of $\mathbb{C}(z)$, we see that if $\int fe^g dz$ is elementary, then we can write

$$ft = \sum_{i=1}^{n} c_i \frac{u_i'}{u_i} + v',$$

with $c_1, \cdots, c_n \in \mathbb{C}$ and $u_1, \cdots, u_n, v \in \mathbb{C}(z,t)$. Now let $F = \mathbb{C}(z)$, so that $f, g \in F$ and $u_1, \cdots, u_n, v \in F(t)$. By factoring each u_i as a power product of irreducible elements of $F[t]$ and using logarithmic derivatives, if necessary, we can guarantee that the u_i's which are not in F are distinct monic irreducible elements of $F[t]$. Imagine v expanded into partial fractions with respect to $F[t]$. The lemma implies immediately that the only possible monic irreducible factor of a denominator in v is t, which is also the only possible u_i not in F. Thus v is of the form $\sum b_j t^j$, for j ranging over some set of integers and each $b_j \in F$. Since $\sum c_i u_i'/u_i \in F$, we have $ft = (b_1' + b_1 g')t$. Writing $b_1 = a$, we have $f = a' + ag'$, with $a \in \mathbb{C}(z)$. Conversely, if there is an $a \in \mathbb{C}(z)$ such that $f = a' + ag'$ then one elementary integral of fe^g is ae^g. Thus fe^g has an elementary integral if and only if there is an $a \in \mathbb{C}(z)$ such that $f = a' + ag'$.

For given $f, g \in \mathbb{C}(z)$, the possibility of finding $a \in \mathbb{C}(z)$ such that $f = a' + ag'$ can be decided by considering partial fraction expansions for f, g, and a. For $\int e^{z^2}dz$ we have the equation $1 = a' + 2za$, which is easily seen to have no solution $a \in \mathbb{C}(z)$. For $\int (e^z/z)dz$, we have the equation $1/z = a' + a$, which also has no solution in $\mathbb{C}(z)$. Therefore $\int e^{z^2}dz$ and $\int (e^z/z)dz$ are not elementary. By certain changes of variable we can get other nonelementary integrals. For example, if we replace z by e^z in the second integral we get $\int e^{e^z}dz$ nonelementary, and replacing z by $\log z$ we get $\int (1/\log z)dz$ nonelementary. The integral $\int \log\log z\,dz$ reduces to the previous integral by integration by parts, so it also is nonelementary.

It is slightly more complicated to show that $\int (\sin z/z)dz$ is not elementary. To do this, first change the variable to $\sqrt{-1}\,z$ to slightly simplify the problem to that of showing that $\int ((e^z - e^{-z})/z)dz$ is not elementary. Here again consider the differential field $\mathbb{C}(z,t)$, where $t = e^z$. If our integral is elementary, Liouville's theorem enables us to write

$$\frac{t^2-1}{tz} = \sum_{i=1}^{n} c_i \frac{u_i'}{u_i} + v',$$

with $c_1, \cdots, c_n \in \mathbb{C}$ and $u_1, \cdots, u_n, v \in \mathbb{C}(z, t)$. Again write $F = \mathbb{C}(z)$, so that $u_1, \cdots, u_n$, $v \in F(t)$, again arrange that the u_i's which are not in F are distinct monic irreducible elements of $F[t]$ and that v is expressed in its partial fraction form, and use the lemma. We again get that the only possible u_i not in F is t, so that $\Sigma c_i u_i'/u_i \in F$, and the only possible monic irreducible factor of a denominator in v is t. Writing $v = \sum b_j t^j$, as before, with each $b_j \in F$, we deduce as before that $1/z = b_1' + b_1$, which is impossible. Therefore $\int (\sin z/z) dz$ is not elementary.

7. The question arises whether for any explicitly given elementary function of the complex variable z it can be decided whether or not the function has an elementary integral, and if so, finding it. It is not difficult to see, using the method of the previous section, that this can be done for any function in $\mathbb{C}(z, e^g)$, where g is any nonconstant element of $\mathbb{C}(z)$, but the general question is not so easy. Hardy's book **[1]** discusses the systematic integration of the kinds of elementary functions that occur in calculus, the main point being that there really *is* a system (contrary to the sometimes expressed opinion that integration in calculus is as much an art as a science), but the book barely broaches the general decision question, which very quickly leads to once intractable questions about points of finite order on abelian varieties over finitely generated ground fields. A solution to this decision problem has recently been announced by Risch **[3]**.

References

1. G. H. Hardy, The Integration of Functions of a Single Variable, 2nd ed., Cambridge Univ. Press, New York, 1916.

2. R. Risch, The problem of integration in finite terms, Trans. Amer. Math. Soc., 139 (1969) 167–189.

3. ———, The solution of the problem of integration in finite terms, Bull. Amer. Math. Soc., 76 (1970) 605–608.

4. J. F. Ritt, Integration in Finite Terms, Columbia Univ. Press, New York, 1948.

5. M. Rosenlicht, Liouville's theorem on functions with elementary integrals, Pacific J. Math., 24 (1968) 153–161.

THE COLLEGE PREPARATION FOR A MATHEMATICIAN IN INDUSTRY

ERWIN H. BAREISS, Argonne National Laboratory

I should like to express my deep appreciation to this association for inviting me to speak on industrial mathematics, a subject which has been ignored for many

Erwin Bareiss received his Ph.D. at the Univ. of Zürich under Rudolf Fueter and Rolph Nevanlinna. He worked with the U.S. Naval Ship Research and Development Center — Washington before joining the Argonne National Laboratory. He also holds a professorship of Computer Science and Engineering Science at Northwestern Univ. He has held Visiting appointments at the Univ. of Maryland, Harvard Univ., was a SIAM Lecturer, and has contributed to hypercomplex function theory, mechanics, numerical analysis, transport theory, and other areas. *Editor.*

years by the mathematical elite of America. It is needless to say that such an invitation required courage on the part of the organizers. In my opinion, it is a sign of progress. Looking back in history, and into the future, I volunteer the prophecy that the separation of mathematics into pure and applied mathematics will appear only as a short interlude. I know some of you will disagree with me, but remember all great mathematicians of the past were non-pure mathematicians: Euclid, Fermat, Leibnitz, the Bernoullis, Euler, Laplace, Gauss, Jacobi, Weierstrass, Hilbert, Georg Birkhoff, von Neumann, and many more, including your own favorite name. Let me say that my great teachers were pure mathematicians, namely Nevanlinna, Fueter, Finsler, and Ahlfors. When I requested to write a dissertation in applied mathematics, the chairman told me bluntly, "Mr. Bareiss, you are too intelligent to write a dissertation in applied mathematics." Thus, my graduate work was done in Galois theory and the theory of functions of a hypercomplex variable. This, of course, happened in Zürich, Switzerland, some twenty years ago; and to this day, I enjoy talking about quaternions and Clifford algebras. But after I had received a fellowship to study in the United States, I acquired an additional degree in applied mathematics and engineering, which provided me with my livelihood. I became an applied mathematician.

What has this story to do with the topic of tonight, "The College Preparation of a Mathematician in Industry"? You may guess that it will have a bearing on my answer at the end of my talk. To see the problem in the proper perspective, I shall present background material and thus subdivide the analysis into three parts:

a. The economic state of the mathematical community at present.

b. A historical review of the role of the mathematician in industry.

c. The psychology of employment in industry.

Although I shall draw conclusions and make recommendations, which may perhaps come as a surprise to some of you, the main purpose of this address is that you continue to look at the problem of mathematical education with an open mind, and that you form your own opinion.

a. *The Economic State:*

As part of its recently completed two-year study, the President's Commission on School Finance tackled the ticklish problem of making the educational system accountable for the money it consumes in ever-increasing doses. Its far-reaching report on accountability is of vital concern to you, the taxpayers, parents, and teachers.

If education is to compete successfully with other increasing demands on public treasuries, its proponents must be able to demonstrate that whatever funds are provided are achieving the desired results. This is extremely difficult because of the intangible nature of its product — learning.

Educators are expected to perform functions which impart to students the knowledge of skills such as the command of language, writing, and mathematics, and they

can, and should, be held accountable for their ability to teach those skills. In addition to skills, they must try to develop for students a desire to learn, an attitude and ability to relate with others. These latter student attributes are not easily measured.

The attempt to determine how well students have learned skills is not new. What is new, and what is now seriously lacking, is the ability to determine how well the student, as an individual, has benefited from his school experience.

Against this background of accountability, i.e., testing of students and teachers, R. D. Anderson, in the *Notices* of AMS, February, 1972, gave a gloomy picture of academic employment prospects for pure mathematicians. According to his figures, of 1300 pure mathematicians seeking a job for September, 1972, there will remain 500 $\pm$ 200 pure mathematicians unemployed. How and where should these mathematicians earn their living? In industry, government, and in the military service. As we all know, the research policies of the federal government have been undergoing some major changes during the last three or four years. The directions of these changes are clearly defined in President Nixon's budget message relating to fiscal 1973. A quick reaction to that part of the message which refers to research and development is that technology and applied research will be funded at a level which reminds one of the affluence of five or six years ago, but basic science will be funded at a level only a little better than at present. Furthermore, Dr. William D. McElroy, NSF director at the time the budget message was put together, is quoted as saying that the nation does not need any more research Ph.D.'s. Actually, NSF foresees a 14.7% surplus of Ph.D. scientists in 1980. The following figures are quoted from Manpower Comments (MC) No. 7, July–August 1971.

In 1980, total supply of Ph.D. scientists will be 325,000, with a surplus of 41,700, i.e., 14.7%. The supply in mathematics will be 24,350 with a surplus of 2800, or 13%. The lowest surplus is forecasted for the Physical Sciences with a supply of 82,250, and a surplus of 400, or $\frac{1}{2}$%, while the highest surplus is seen in engineering, with a supply of 55,650 and a surplus of 16,100, or 40%.

From a national point of view, the unemployment rate of scientists in 1971 is not alarming. According to a survey published in MC 8 (1971), based on a sample of 253,078 respondents, the average unemployment rate was only 2.6%. Of the 19,745 mathematicians of this sample, 491, or 2.6% (i.e., the average rate) were unemployed, while of the 8840 computer scientists 309, or 3.6% were unemployed. Higher unemployment rates are recorded in sociology (3.8%), physics (3.9%), and linguistics (4.5%).

So far, we have been concerned only with Ph.D.'s in mathematics. The U.S. Department of Labor, Bureau of Labor Statistics, estimated the total number of all technicians employed in 1970 at 1,010,400. Of these, only 5800, or .6%, are classified under mathematics. From a purely economic point of view, these figures are insignificant. However, it is interesting to note that the anticipated need for these mathematics technicians for 1980 is 10,100, or an increase of 74.1%, far beyond all other classifications and almost twice the 38.1% anticipated average increase needed for all technici-

ans. Unfortunately, computer programmers and assistants to scientists, engineers, and surveyors are summarized under one category, which included 143,600 technical employees (or $\sim$ 14%) in 1970. For this category, a need of 206,000, or an increase of 43.3 %, is expected for 1980.

Now, it is interesting what Frederick E. Terman, Vice-President and Provost Emeritus at Stanford University has to say on "Supply of Scientific and Engineering Manpower: Surplus or Shortage" (Science, July 1971).

After many years during which Ph.D.'s appeared to be in short supply, people with new Ph.D. 's in certain areas are having difficulty in locating satisfactory jobs. The problem is not one of unemployment. Rather, the problem lies in the inability to satisfy the new Ph.D.'s job expectations, after having been led by teachers and advisers to believe that investing time and money in the Ph.D. would be the key to an exciting and attractive career. The disillusionment is greatest for those students who studied at the most prestigious schools, because their expectations were the highest.

During most of the 1960's, government-supported research and expenditures for defense and space work were rapidly growing. As a consequence, the growing number of Ph.D.'s produced each year was readily absorbed until the 1969–70 academic year.

First, the number of openings for young Ph.D.'s at universities suddenly and sharply dropped. This was partly because, beginning in 1969–70, the number of students studying science and engineering abruptly leveled off. Concurrently, government funds for academic research leveled off, so that research and associated graduate activities in universities stopped expanding; in some cases they decreased.

The total number of men graduating annually in science and engineering has grown greatly since 1955, although it has tended to level off since 1960. This is a situation with which observers of the educational scene are familiar. However, what has not previously been recognized is the fact that a retreat from science and engineering among men began much earlier than is generally assumed, when expressed as percentage of total enrollment. The graduating classes of 1962, which signaled the start of the decline, were in high school at the time of Sputnik. Thus, if Sputnik had any effect on American youth's interest in a career in science or engineering, the effect was negative. Interest in the biological and mathematical sciences has increased during the last decade, at the expense of engineering and the physical sciences.

A special situation exists in mathematics, since the student who has received a B.S. in pure mathematics has little that is marketable in terms of employment. However, the student who holds an M.S. in traditional mathematics is qualified to teach in a high school, in a community college, or in a liberal arts college that cannot attract and hold Ph.D. 's. In addition, industry seeks people with an M.S. in mathematics to work in such fields as statistics, continuum mechanics, biomathematics, operations research, and, particularly, computer science. The number of M.S. degrees in the mathematical sciences awarded annually to men was 1428 in 1960, but 4202 in 1968, an annual average increase of 14%.

To sum up the situation for Ph. D.'s, we quote from a recent issue of the *Notices* of AMS signed by Richard D. Anderson (Louisiana State University), William L. Duren, Jr. (Chm., University of Virginia), Gail S. Young (University of Rochester), and Dr. C. Russell Phelps of the Conference Board of the Mathematical Sciences:

Not only industry but colleges as well are now seeking applied mathematicians in preference to the pure mathematicians. There are about 1000 jobs in four-year colleges and universities where a Ph.D. will be needed; colleges seek or prefer pure mathematicians for less than half of these positions (according to a current CBMS survey). With Computer Science and Applied Mathematics programs coming on stream, the pure mathematicians can hardly expect to be employed in industry in large numbers.

b. *Historical Review:*

The history of mathematics in industry goes back some 80 years. In 1888, a young German mathematician had his Ph.D. dissertation accepted, but he never received the degree. He was a socialist agitator whom the police decided to arrest. In a story-book fashion, he fled in the middle of the night to Switzerland. The following year he came to America, and joined what is now the General Electric Company. According to a scholarly study by C. T. Fry (Science 1964), he was the first mathematician employed in industry. The man was Charles Proteus Steinmetz, and the title of his thesis sounds as modern as 1972: "On Involutory Self-Reciprocal Correspondences in Space which are Defined by a Three-Dimensional Linear System of Surfaces of the n-th Order." Steinmetz was a charter member of the American Mathematical Society and participated actively in its affairs.

Prior to this time, industry had been flourishing through inventions of the purely Edisonian type. But the problem of transmission in telephony, and the problems of transmission and generation in the power industry, raised questions of a more subtle and analytical type and required a more scientific approach. Industrial research was born.

In a report to the National Resources Planning Board, published in 1940, C. T. Fry from the Bell Telephone Laboratories made a serious attempt to estimate the number of professional mathematicians working in industry and came up with the figure 150.

This, of course, involved a matter of definition. In 1940, as today, many industrial physicists, chemists, and engineers had considerable mathematical training and ability and were using it in their work. It would have been foolish to count all these as mathematicians. Fry resolved the difficulty by counting the members of the American Mathematical Society who clearly indicated industrial or government employment. He argued that a scientist who had sufficient interest to belong to a society devoted exclusively to creative mathematical research could properly be defined as a mathematician.

This study was made in 1939, half a century after 1889. It might be interesting to fill in the quarter-century points.

The membership list of 1964, by a sampling process, gave a count of 1800. For 1914, depending on whether or not some doubtful cases are included, one gets 11 or 15, say $12.2 = \sqrt{150}$.

Using these figures, and recording a "1" for Steinmetz in 1889, 12.2 for the year 1914, 150 for 1939, 1800 for 1964, we record with amazing consistency an exponential growth of industrial mathematics which gives a 12.2-fold increase every 25 years, or an annual average increase of $11\frac{1}{2}\%$. By extrapolation, we obtain

$$\begin{array}{rcl} 4{,}000 & \text{for} & 1972 \\ 22{,}000 & \text{for} & 1989 \\ 270{,}000 & \text{for} & 2014 \end{array}$$

and

$$40{,}000{,}000 \quad \text{for} \quad 2064.$$

Of course, these predictions are made in jest and for a hearty laugh! However, the 1972 prediction is realistic and demonstrates the large number of mathematicians in industrial and governmental laboratories, an environment which only a few generations ago would have been judged inhospitable.

I made some phone calls to a number of industrial laboratories, such as the IBM Research Center in Yorktown Heights, New York and the brand new Bell Telephone laboratory in Naperville, near Chicago, and found out to my surprise that such Laboratories have on their payrolls more than 10% college trained mathematicians based on total employment. However, in industrial as opposed to government Laboratories, the job title mathematician does not exist. The mathematicians are classified as scientists or communication engineers, etc.

In industry, the first period (1890 to 1915) can be characterized as one of handbook engineering. It is difficult for us to appreciate how primitive engineering was. One of the things upon which Steinmetz' fame rests was his success in training electrical engineers in the use of complex quantities in alternating-current theory. Mathematicians had been using the method for decades. But the vast majority of electrical engineers found it incomprehensible, and were completely mystified that the square root of minus 1 should have anything to do with electric currents. But, with the new century, things began to happen in physics, then in chemistry, and in the end in engineering and the society as well. The growth in employment of mathematicians in industry is one aspect of this revolution.

The quantum hypothesis was formulated in 1901. The vacuum tube was invented in 1907. The special theory of relativity was published in 1908. Bohr's paper on the hydrogen atom appeared in 1913, and Mosley's on atomic numbers in 1914.

The period from 1915 to 1940 was equally exciting, though in a quite different way. It was the period of quantum mechanics and electronic physics. Chemistry moved

explosively ahead during this period under the impetus of the clear-cut structural ideas which grew out of the work of Bohr, Mosley, and the Braggs.

It was also a period of tremendous change in industry, which discovered that profits could be derived from scientific research, as distinguished from engineering developments. Research laboratories sprang up by the hundreds, many in industries in which management was ill-equipped to direct them or even to understand the nature of their activities. Scientific research was now being consciously organized and exploited by industry.

The practical engineer got his mathematics mostly through self-education. He did not question its value. Conscious of his own limitations, the engineer tended to give a high rating to anyone with mathematical training and interests who was reasonably articulate, regardless of his true mathematical ability. A talented mathematician who attempted to cooperate with his engineering associates was rewarded with respect and appreciation. To prove this statement we note that between 1931 and 1933, the depression years, the professional staff of Bell Telephone Laboratories was reduced by about one-third, but not a single member of the Mathematical Research Group was dismissed.

While the goal of industry is to make a profit, the "true" mathematician is only concerned with ideas. But there are many mathematicians who are deeply interested in both ideas and things. Hence a good mathematician may also be a good engineer. In an industrial environment there is a strong tendency to assign such a man the duties and responsibilities of an engineer. However, when this is done, the mathematicians remaining available for consultation are those who are only interested in ideas and who for that reason may be the least effective consultants.

The period from 1940 to 1965 can be called the period of particle physics. Of course, there have been important advances in other areas, but none had the social impact of controlled and uncontrolled nuclear power. Information theory, which in effect quantizes all intelligible thought, also made an impact on the scientific scene, and may lead to social changes not now foreseeable.

And, finally, there is control theory. The electronic computer is an omnipresent realization of it. We now have the ability to control systems of all kinds, from the simplest machine to the most involved spacecraft, not through rigid procedures but through flexible processes akin to thought, where the only invariant is the underlying system of logic. This is so important that the world will never be the same again. It may well be that 50 years from now computer application will stand out as the great scientific achievement of the period.

In the industrial research laboratory, the most important evolution has been the teamwork. Without the team approach we could not have effectively exploited the new materials and new theories in mathematical sciences. The team approach expanded the limitations of an individual brain by linking several or many brains into a single interacting system—a system which is as necessary for the final accomplishment as are the materials or the scientific theories.

With the emergence of the team and to some extent because of it, the place of the mathematician in industry has become more complex and perhaps more central. To understand this, we shall take a brief look at the nature of mathematics.

What is mathematics? The best definition I have heard is probably this one: "Mathematics is what mathematicians do." More down to earth, mathematics is an *art*, a *language*, a *tool*, and a *means of accounting*.

As an art, it deals with postulates and their logical consequences. It is creative and has no necessary connection with the physical world.

To create his work, the artist uses the language of mathematics. But conversely, the mere use of the language of mathematics does not imply creative art work. For example, many physical laws can be stated by means of equations, i.e., in mathematical language.

Now, it may become possible to make use of known mathematical theories and arrive at the physical consequences of the laws. To state it more simply, we often can solve the difference, differential and integral equations which are the mathematical models of physical processes and thereby derive formulas or assign numerical values for the electrical current, neutron flux, wave motion, plastic flow, or population growth. In doing so, we're using mathematics as a tool.

From the point of view of the professional mathematician, the art of mathematics ranks as most prestigious, followed by language and tool. Accounting is often not considered mathematics at all. From the standpoint of industry, the order of importance is reversed, for the art has often no necessary connection with the physical world and is therefore of little immediate value, whereas the language and the tool clearly have value, and, without accounting our techno-society could not exist.

It seems that the period since 1965, the last quarter of the first century of industrial and planned research, must be characterized as the period of reckoning, or more specifically, the period of accounting. Management and Congress are no longer impressed by the sheer sight of mathematical language. The slogan is now: Accountability. And in spite of the tremendous progress the sciences, and in particular the mathematical sciences, have made, the general attitude is one of disappointment, almost mistrust. In the last few years I had to review quite a number of research proposals whose only attribute was an impressive mathematical language without much deep thought, and of little mathematically aesthetic or even practical value. But they sounded impressive! Of course, no sincere mathematician deliberately tries to sell substandard merchandise in a pretty package to the disadvantage of an honestly good proposal in a simple language. As always, there are two sides to every story, and I do not excuse administrators and management from their responsibility.

I hope that the fourth quarter will not remain a period of consolidation and reckoning only, but a period of new awakening and of distinct progress in mathematics. I foresee a regrouping of values in industrial mathematics. Only recently the first patent was granted for a computer code. It may well be that in the near future patents

will be granted for new methods or algorithms to solve computational problems. And before the end of the first century of industrial mathematical research, the federal patent law may be written so that patents can be granted for new mathematical theorems in recognition of the economic value of creative mathematics.

c. *Psychology of Industrial Employment*

It is interesting to note that at the height of scientific glory in 1968, a book appeared with a title comparing the scientists with the highest caste in Hinduism. The title of the book is "The New Brahmins, Scientific Life in America" by Spencer Klaw. It is well written, and was certainly designed to become a best seller such as Caplow and McGee's "The Academic Market Place." Did any one in the audience read this book? Not many. The book appeared just at the time when the wholesale firing of scientists began. In spite of the untimely title, the author has taken a realistic view of science and mathematics. He notes:

In the United States, the marriage of science and the practical arts that Robert Hooke hoped the Royal Society would bring about has been consummated mainly in the laboratories of large corporations. The industrial laboratory, rather than the university, is now the principal habitat of the scientifically trained American.

Employees of the Bell Laboratories have made many important contributions to both science and technology besides the transistor. These include the mathematical theory of information, formulated by Claude Shannon, which led to major improvements in the coding, transmitting, and switching of messages. But despite the enormous benefits that American Telephone and Telegraph and certain other companies have gained by hiring good scientists and giving them their heads, the number of scientists in industry who are free to do [independent] research is relatively small. Only a large company with a fairly stable business and good profit margins—or a regulated monopoly like A.T.&T., which can charge the cost of its research to its customers — can afford to invest large sums of money in undirected basic research.

Many scientists fail to win autonomy and become problem solvers. They are assigned to groups engaged in what may be described as exploratory development. Such groups are not expected to explore unknown scientific territory; they are charged, rather, with finding the best route across scientific terrain whose main features are well known but have not yet been accurately mapped. Most of this work is done by engineers, and by scientists who have only bachelor's or master's degrees, who together constitute the great majority of all professional workers in industrial research and development. But a fair number of scientists with Ph.D.'s are also involved, not only as leaders of groups but as members of the rank and file. Some become part of the proletariat of industrial research, carrying out routine tasks under fairly close supervision.

Harvey Sherman, former president of the American Society for Public Administration, observed that the typical corporate executive sees the scientist as a "narrow

specialist with no interest in efficiency or economy or in the overall objectives of the enterprise, a person who...objects to all types of control, and who is more interested in impressing other members of his profession than in the success of the enterprise for which he works." Sherman noted that the scientist takes an equally dim view of the executive: "By and large, the scientist sees [him] as a bureaucrat, paper shuffler, and parasite; an uncreative and unoriginal hack who serves as an obstacle in the way of creative people trying to do a job, and a person more interested in dollars and power than in knowledge and innovation."

The Harvard Business School professors Ralph M. Hower and Charles D. Orth III reported in their book *Managers and Scientists* that the managers of a number of industrial laboratories shared almost universally this point of view of the mid–1950's: "A good man should be promoted to managerial positions; a scientist who rejects an opportunity for such advancement will be held down in status and pay...Indeed, it appeared to us in some instances that men who insisted on staying in research were subject to treatment which in effect constituted punishment."

Industrial research managers often complain about how seldom the scientists who work for them come up with bright and original ideas. This is only partly accounted for by the fact that very bright and original young scientists usually do not take jobs in industry. It is also clear that in most industrial research organizations the climate is unfavorable to ideas that are daring or radical. Change upsets business organizations, and is bound to be strenuously resisted. A former head of research at General Motors once observed: "The greatest durability contest in the world is getting a new idea into any factory."

A scientist who has a good (but radical) idea may have to choose between forgetting about it and risking his job in order to prove its feasibility. Arthur K. Watson, while President of the IBM World Trade Corporation, told an audience of accountants: "The disk memory unit, the heart of today's random access computer, is not the logical outcome of a decision made by IBM management. It was developed in one of our laboratories as a bootleg project — over the stern warning from management that the project had to be dropped because of budget difficulties. A handful of men ignored the warning. They broke the rules. They risked their jobs to work on a project they believed in."

Klaw asserts that originality and imagination in industrial research are also discouraged by the fact that the way a scientist in industry thinks is less important, by and large, than how he behaves. There are scientists who do manage to get ahead purely by exercising their intellectual prowess. But money, freedom and power are more commonly won by exercising nonintellectual skills of the kind that are rewarded in other walks of corporate life. To begin with, salesmanship counts heavily. When a scientist in industry suggests a particular line of investigation, or a particular attack on a problem, acceptance of his proposal may depend less on its intrinsic merit than on his ability to convince other people, who are not scientists, of its commercial or technological relevance. This skill is perhaps most highly valued in laboratories that do a

great deal of research under government contract. Senior staff people at such laboratories spend a lot of time writing up proposals for new projects and trying to persuade prospective clients to support them. "We put a lot of emphasis on communication, both oral and written," said the personnel manager of a laboratory that works mainly for the National Aeronautics and Space Administration. "I wouldn't care if you could guarantee me [that] a man is a genius, I wouldn't hire him if he's not articulate."

Unfortunately, laboratory administrators themselves are probably more often picked for persuasiveness than for brains, and may have a lot of difficulty themselves in telling good ideas from bad.

Salesmanship is not the only nonintellectual talent that pays off in industrial research. Young scientists are also given high marks for tact, dependability, and the ability to work smoothly with other people. Some of the biggest (and best) laboratories do tolerate a certain number of oddballs who like to work at night, or who are incapable of meeting deadlines, or who refuse to tell their supervisors what they are up to. Often, however, their position is valued by the Laboratory's management mainly because their presence proves that what the laboratory really cares about is not sterile conformity, but creativity. Many laboratories take great pains to screen out scientists with the wrong kind of personalities. In particular, the neophyte must take care not to seem brash or overeager. The author of "Introduction of the Newly Graduated Scientist to Industrial Research," (Research Management, 1960) and staff specialist of a large oil company, emphasizes how important it is for a new recruit to be modest. Then he adds: "Still another problem is that of the impact of corporate policies, ways of doing things, communication channels, etc., upon the neophyte scientist who is at the stage of his life where he is properly most eager to accomplish great things. He may soon discover that his earnest and well-intentioned efforts may have earned him the unofficial, yet damning title of 'boat rocker'."

Educational requirements

I shall now turn to the last part of my talk, the Educational Requirements for a mathematician in industry. You may have already made up your mind, but let me quote the opinion of three different sources, and then add my own conclusions.

T. C. Fry, former director of the Mathematical Research Department at Bell Telephone Laboratories is quoted from *Science* (Vol. 143):

"To be an effective member of [a] team, the mathematician must also understand the basic principles of the various disciplines which he is expected to discuss. He should be, in other words, the sort of man who a century ago was known as a natural philosopher — a man who had a keen analytical mind, adequate mathematical training, and a broad and sympathetic interest in a wide range of natural phenomena. There is already a clear need for such men, and, in my opinion, this may well become the most important role the industrial mathematician of the next generation will play. If this judgment is correct, we may well ask where these men are to come from.

"Those I have known have often been physics or engineering undergraduates who developed a love for mathematics and majored in it for their doctor's degrees. This was true, for instance of Bode, MacMillan, Schelkunoff, and Shannon... .This is not hard to understand, since such men have interest both in ideas for their own sake and things.

"But while this is an effective pattern of education, the reverse — an undergraduate major in mathematics followed by a Ph.D. in science — does not have equivalent value. The reason is that the ingredient which the mathematician adds to the team is his greater emphasis on precise definition of terms and rigorous logical analysis, an emphasis seldom obtained outside the graduate mathematics curriculum.

"There is, then, a legitimate need for graduate mathematical training which is both sound mathematically and sympathetic to the phenomena of the real world. Whether we call it applied mathematics or something else makes little difference. Its object is to train men who can be natural philosophers. [This requirement runs exactly counter to the goals of our best mathematics departments in the country.

"We need, I think, in the universities and the mathematical society as well, a broader concept of the social value of mathematics. Not a de-emphasis of the art, for that would be a tragedy, but a greater pride in the full scope of the discipline and a stronger interest in its social values. Such a concept would greatly facilitate the training of the "natural philosophers" which industry will increasingly need in the foreseeable years ahead."

A week ago I received a memo from a senior member of a national laboratory whom I had always ranked as loyal defender of pure mathematics. This memo reads in part:

"...we have gradually come to the conclusion that some kind of an explicit apprentice or intern system could be very useful in allowing the [department] to engage effectively in consulting activities. The argument is based on the following two points:

1. Persons capable of consulting work in applied mathematics are best trained by actual experience in problem formulation and solution, preceded of course by adequate academic preparation.

2. This experience is best obtained by initial work in association with persons who already have experience in consulting activities, somewhat in the same manner as is done in other professions such as medicine and law."

These two points are followed by the remark:

"We incline toward the recruiting of engineering graduates for this work rather than mathematics majors, although exceptions should certainly be made in individual cases."

Anderson, Duren, Young, and Phelps have this to say in the *Notices* of AMS mentioned earlier:

"...In any case, there is no reason to make a headlong switch to applied mathematics. The academic outlet for applied mathematicians is limited, and we don't know what kind of applied mathematics will be needed. Right now, it is the computer

scientists who are in demand, but very soon the environmental and other human problems may require more mathematicians of operations research type. These prospects change too fast to serve as a basis for long-term educational programs. Perhaps what is needed is sound graduate education in [pure] mathematics with provision for continuing education, both before and after the doctorate, in the applications of mathematics. Nothing definite can be said at this time."

Let me now summarize my opinion:

The training of a mathematician for government, industry, or insurance depends on the educational level (B.S., M.A., or Ph.D.) at which he wants to leave school.

If he plans to earn only a B.S. degree, a solid education in a computer science department is most desirable. Should such a department not exist at his college, I urge strongly that the following courses be given in the mathematics department:

Introduction to Computer Programming
Classical and Modern Linear Algebra
Theory of Computation and Numerical Analysis
Probability and Statistics.

If he plans to leave school with an M.S. degree, I consider the above-mentioned courses also as an absolute must. The remaining courses should be tailored such as to develop the student's strongest sides. At least one course must be taken which requires absolutely rigorous logical deductions.

For Ph.D. candidates, an undergraduate degree in mathematics is not necessary because a rigorous mathematical training can be obtained in graduate school. It is harder to acquire familiarity with physics, chemistry, or computer science on a postdoctoral level.

However, the best preparation for a mathematician in industry comes through the attitude of the student's teacher toward applied mathematics, especially when the student has also been taught to teach himself the facts and the meaning of the ever-changing world of computational and applied mathematics.

To be successful in industry, additional qualities are needed as demonstrated above. These are considered beyoned the responsibilities of the mathematics departments.

Is this last sentence a true statement?

You form your own opinion.

Presented to the Rocky Mountain Section, Southern Colorado State College, Pueblo, on May 5, 1972.

THE MATHEMATICAL SOCIETIES AND ASSOCIATIONS IN THE UNITED KINGDOM

THOMAS WILLMORE, University of Durham, England

The Mathematical Societies and Associations in the United Kingdom fall roughly into two classes — those primarily concerned with mathematical research and those concerned essentially with the teaching of mathematics.

The Royal Society of London is the oldest and the most respected scientific learned society in the U. K., but this is not concerned solely with mathematics. Moreover, although its influence on contemporary mathematics is substantial, its membership is restricted to a very small number, known as Fellows of the Royal Society. The Royal Society is primarily concerned with research, though it is represented on many important committees concerned with the teaching of mathematics.

The London Mathematical Society is the British equivalent of the American Mathematical Society. Again it is primarily concerned with mathematics research, and although during the last two decades it has sponsored and arranged instructional conferences, these are essentially at the graduate level. Its membership is drawn essentially from university teachers, teachers in polytechnics, and some professional mathematicians in industry. Its attitude is essentially academic in outlook and, although it publishes papers in applied mathematics, the majority of the papers in its journals are concerned with pure mathematics.

After the 1939–46 war, many British mathematicians regarded the London Mathematical Society as too traditionally oriented to deal with the explosion of mathematical research. In an attempt to provide a shot in the arm, the British Mathematical Colloquium was born, and held its first meeting in 1949 at Manchester. This was conceived as an annual conference of British mathematicians to be held in turn at different universities — the idea was to attract world acclaimed mathematicians to give hourly addresses and follow them by short papers, splinter groups — specially to encourage younger research mathematicians to let their work reach a wider audience. Relations between the London Mathematical Society and the British Mathematical Colloquium are extremely harmonious, and each plays an important complementary role.

In 1968 the Institute of Mathematics and its Applications was born, due primarily to the energy of Professor M. J. Lighthill. This was intended as an institution for the professional mathematician in industry, as well as academic mathematicians. Although this institute has been in existence for only a few years, it has already amply justified its existence. It publishes a journal and a bulletin and is primarily oriented towards applications of mathematics.

The Mathematical Association, founded in 1870 with the immediate objective

of improving the teaching of geometry in schools, is perhaps the nearest equivalent to the Mathematical Association of America. Its membership consists largely of school teachers, though it has many members from industry, colleges of education, polytechnics, and the universities. Recently the Association celebrated its Centenary Meeting in London, April 1971. On this occasion we were delighted to receive the congratulations of the Mathematical Association of America, represented in person by Professor Harley Flanders. The American Association presented to its British counterpart a certificate commemorating the achievement of a hundred years of successful activities.

The *Mathematical Gazette* of the Mathematical Association is well known and has contributed much towards the development of mathematical education throughout the world. Until recently it was the main British journal which carried reviews of new mathematical books — now that task is shared between the *Gazette* and the *Journal of the London Mathematical Society.*

However, it could be argued that its main contribution to the study of mathematical education is to be found in the specialised reports, published by the Association. These are over sixty in number and the topics dealt with vary from report No. 56 "Applications of Sixth Form Mathematics" (1967) to No. 61 "Primary Mathematics — A Further Report" (1970).

Much has been written about the successes of the Mathematical Association — perhaps the best account was given by Professor M. J. Lighthill in his Presidential Address to the Centenary Meeting of the Mathematical Association in London, April 14, 1971. This address was published in the *Mathematical Gazette,* June 1971, pp. 249–270. If in this note I dwell more on its deficiencies this by no means implies that I do not agree with Professor Lighthill's justified comments on its successes.

My main criticism is that the Mathematical Association seems to rely for its support on the middle aged or the "more than middle aged". A casual glance at those attending the Annual Conference in London 1971, showed that there was a very poor attendance from the under thirties. Of course, it is expensive to attend a conference in London and many teachers were unable to obtain contributions from their local education authorities — one of the mysteries of the British educational system is that "instructional courses" qualify for such grants but "annual conferences" which may be educationally more valuable may not. The necessity for self-financing would naturally fall more heavily on the younger members of the association, and this may partly explain their absence. I feel, however, that the main reason for lack of support from the younger teachers is because they consider the approach of the Mathematical Association too traditional. The startling changes in teaching methods which have taken place in the primary schools in the U. K. during the last two decades have not reached the senior levels of many secondary schools, which still proceed along traditional lines. Many younger teachers feel that the Association of Teachers of Mathematics (A. T. M.) is a more lively organisation which is more concerned

with the practical difficulties arising in the classroom. It may be of interest to sketch the beginnings of the A. T. M.

The A. T. M. was founded, as the Association for Teaching Aids in Mathematics, in 1952, by R. H. Collins (who had earlier launched 'Mathematical Pie') and C. Gattegno (who was then a mathematics methods tutor at the London Institute of Education). Soon afterwards, Cyril Hope and Trevor Fletcher joined it. These four, and particularly Gattegno, gave the Association its flavour in those years of the early fifties.

Although it was the weekend seminars organised privately by Gattegno about the teaching of mathematics that created the actual occasion for the launching (since he met the founder members at these meetings), it is clear looking back that the time was particularly appropriate for starting a new Association. In no special order, there were these influences: the relatively new 'secondary modern' schools, established as a result of the 1944 Act, were teaching mathematics to a wide band of ability and meeting problems in their teaching which were not solved by the usual 'tips and wrinkles' type of advice that was available; the Mathematical Association had not yet turned its attention to these problems, and the *Gazette* seemed very distant from the classroom; a number of teachers had already begun to experiment with the use of 'teaching aids' in order to make mathematics more accessible; there were a number of significant activities on the continent — Piaget's work was just getting known, with its outline of a theory of concrete activities leading to abstract understanding; a Swiss, J. L. Nicolet, had shown how animated films could teach geometry; several French mathematicians influenced by Bourbaki were beginning to think of the applications of this work to school level. Gattegno brought a wide range of international contacts through which these latter influences worked and offered the chance of breaking the insularity of British theories and methods. No one else was doing this at the time.

But probably the strongest strain that Gattegno contributed was his own intellectually tough form of pedagogy, quite different from the rather esoteric stuff offered in training courses. He offered an ideal of a 'learner-centered' education which began from the premise that teaching should release the energies and abilities of children, which most 'traditional' teaching inhibited except with the very ablest pupils. He also opened up the question of research into mathematical education by asserting that the classroom was the laboratory, and that every teacher could, by his personal research in his own classroom, contribute to the improvement of mathematics teaching.

Although this is an idealised portrait, the A. T. M. has certainly retained over the years (a) a number of international contacts, (b) a concern with the whole age and ability ranges, (c) an interest in films and other aids to teaching, (d) a plea for sensible modernising of the curriculum, and (e) a faith in children's ability to investigate and explore mathematics for themselves.

After publishing on duplicated paper a few sporadic issues of a bulletin, *Mathematics Teaching* was started in 1955 under the editorship of Fletcher.

In 1962 the Association changed its name to the present one on the grounds that its interests were wider than the old name indicated. Also in 1962 a group of members collaborated in writing *Some Lessons in Mathematics* (published by C. U. P.) which became a best-seller in the cause of modernising school syllabuses. The membership was about 500 in 1958, 1000 in 1960, 3000 in 1964 and currently upwards of 6000 (over 1000 are overseas members).

The challenge from the A. T. M. has had a beneficial effect on the Mathematical Association. By the publication of a new journal by the Mathematical Association, *Mathematics in School* (Vol. 1 No. 1 November 1971) that association frankly admits that some new approach to the teaching of mathematics is necessary and that articles in the *Gazette* on mathematics teaching which have appeared during the last hundred years are inadequate.

The Mathematical Association appears to have taken over Gattegno's philosophy that in mathematical education the classroom is the laboratory. All success in its new venture.

Some have argued about the desirability of combining the two associations, namely the Mathematical Association and the Association for the Teaching of Mathematics. The idea certainly has attractions. However, two distinct organisations have the advantage of influencing one another — it is reasonable to assume that *Mathematics in School* was stimulated by the success of the A. T. M. Friendly rivalry promotes competition and forces each organisation to react to new situations. I think it would be a tragedy for the Mathematical Association to become moribund, and this seemed a possibility to me. However, the new periodical should give it another lease of life. Unless the organisation can recruit young active members, it will surely cease to be effective in the future.

The lack of liaison between the London Mathematical Society, the Mathematical Association and the Association for the Teaching of Mathematics was clearly shown by the Centenary Meeting in April 1971. On Thursday, April 20th, there occurred *simultaneously*

(i) the Annual General Meeting of the Mathematical Association in London,

(ii) the Annual Conference of the A. T. M. in Southport, Lancashire,

(iii) a special meeting of the London Mathematical Society at Imperial College, London.

Perhaps overlap with (iii) was unfortunate. But overlap of (i) and (ii) shows a serious lack of foresight for which at least one organisation must take blame.

In conclusion I must thank David Wheeler, Editor of *Mathematics Teaching* for useful information about the beginnings of A. T. M. Needless to say, although I am a member of the Mathematical Association, the Association of Teachers of

Mathematics, and the London Mathematical Society, the views given in this note are my own and should not be interpreted as the viewpoint of any one of these organisations. Long live all organisations concerned with the teaching of mathematics!!!

A LOOK AT THAT 1971 MAA INFORMATION SERVICES SURVEY

L. H. LANGE, California State University, San Jose

1. Introduction. At midyear in 1971, President Victor Klee appointed an *ad hoc* Committee on a Survey of the Membership of the Association. (The Committee: E. F. Beckenbach, Chairman; D. L. Bernstein, J. Hashisaki, L. H. Lange, K. O. May, I. Niven, A. Rosenberg, A. B. Willcox.) The survey was to involve questions whose answers could be helpful in making decisions concerning revised journal privileges and options to be offered to Association members. More generally, the idea was to ask questions related to the information needs of our members. The list of questions was expanded and, for example, it came to include questions about the timing of our national meetings—since many of our schools have drastically changed their schedules in recent years. This, then, is a brief report on the survey which developed and on some of the resulting actions of a responsive leadership which even now are beginning to flow from it.

2. The questionnaire and a tabulation of responses. The questionnaire reproduced below was mailed to the 18,311 members in October. (There were 17,899 domestic members and 412 foreign members.) By the November 15 deadline, approximately 6000 responses were in. Then, by January 1, 1972, 6748 responses had been received in the Washington office—a gratifying volume of response which exceeded by far all of our guesses made earlier. (One correspondent did wonder whether requiring survey respondents to pay 8c postage might not have introduced a bias in the survey. He asked, "Are the less enthusiastic members as likely to think it's worth the postage?") (An Australian member, whose questionnaire reached her by surface mail on December 12, sent the advice that we'd better use airmail in the future if we want on-time responses from down there.) The tabulations below involve those 6748 responses. An earlier look at the first 728 responses received seems to indicate that a restriction of the mailing to a sample of the membership might have served our purposes—though, of course, there would then have been fewer members who would have had the chance to forward their comments and suggestions (as they were invited to do at the end of the questionnaire).

Here is the questionnaire, in toto, along with the numbers and percentages which show the distribution of the replies received. For example, 1832 respondents checked the reply numbered 5.4, and 1832 is 27.1% of 6748.

MAA INFORMATION SERVICES SURVEY

(Slightly altered to fit into the MONTHLY)

			REPLIES	%
1.	(Address information)			
2.	Check one of the following which describes your present principal occupation.	2.1 Employed in a university offering a Ph. D. degree in Math.	1647	24.4
		2.2 Employed in a four-year college or university not offering a Ph. D. degree in Mathematics	2423	35.9
		2.3 Employed in a two-year college	487	7.2
		2.4 Employed in a secondary school	339	5.0
		2.5 Employed in industry	529	7.8
		2.6 Employed in government	296	4.4
		2.7 Full-time graduate student	483	7.2
		2.8 Undergraduate student	73	1.1
		2.9 Other (explain)	404	6.0
		N. R.	67	1.0
3.	Is a subscription to the MONTHLY a dominant factor in your decision to be and remain a member of the Association?	3.1 Yes	4700	69.7
		3.2 No	1994	29.5
		N. R.	54	0.8
4.	Is a desire to support the Association in its efforts to improve the content and teaching of undergraduate mathematics a dominant factor in your decision to be and remain a member of the Association?	4.1 Yes	5083	75.4
		4.2 No	1494	22.1
		N. R.	171	2.5
5.	Check each of the following additional societies of which you are currently a member.	5.1 The American Mathematical Society	3255	48.2
		5.2 The Association for Computing Machinery	532	7.9
		5.3 The Institute of Mathematical Statistics	235	3.5
		5.4 The National Council of Teachers of Mathematics	1832	27.1
		5.5 The Society of Industrial and Applied Mathematics	731	10.8
		N. R.	1704	25.3
6. & 7.	The Association has the option of taking over the TWO-YEAR COLLEGE MATHEMATICS JOURNAL (TYCMJ) effective in 1975. Should the Association exercise this option?	6.1 Yes	2698	40.0
		6.2 No	1000	14.8
		N. R.	3050	45.2

	Have you seen a copy of the TYCMJ?	7.1 Yes	1030	15.3
		7.2 No	5566	82.5
		N. R.	152	2.2
8.	Journal privileges of members include a subscription to the MONTHLY as part of dues, with the option to buy the MATHEMATICS MAGAZINE at a reduced rate. Should the membership be offered various options for subscribing to the MONTHLY, MATH MAG, and TYCMJ?	8.1 Yes	5425	80.4
		8.2 No	704	10.4
		N.R.	619	2.3
9.	By 1975, I would be interested in receiving, with appropriate adjustment of dues, the following journals (check one or more):	9.1 MONTHLY	5624	83.3
		9.2 MATHEMATICS MAGAZINE	2731	40.5
		9.3 TYCMJ	1740	25.8
		N. R.	689	10.2
10.	How much of the MONTHLY do you read?	10.1 Essentially all of it	960	14.2
		10.2 Some, but far from all	4936	73.2
		10.3 Very little or none	812	12.0
		N. R.	40	0.6
11.	For you, are the mathematical articles in the MONTHLY at too high or too low a level?	11.1 Too high	1891	28.0
		11.2 About right	4529	67.1
		11.3 Too low	135	2.0
		N. R.	193	2.9
12.	Effective in 1973, the Summer Meeting of the Association will start on Monday, two weeks prior to Labor Day. At present, the meeting begins one week before Labor Day. Convenient times for me for the Summer Meeting are (check one or more):	12.1 on week before Labor Day	2152	31.9
		12.2 two weeks before Labor Day	2910	43.1
		12.3 three weeks before Labor Day	1715	25.4
		N. R.	1490	22.1
13.	For the Annual (winter) Meeting of the Association, my preference between a meeting some time in January and one between Christmas and New-Year's Day is for a meeting	13.1 some time in January	2965	44.0
		13.2 between Christmas and New Year's Day	1514	22.4
		13.3 no preference	1674	24.8
		N. R.	595	8.8

14. If the Annual Meeting of the Association is to be held some time in January (it is presently scheduled for the third week in January), convenient times for me are (check one or more):	14.1 the first week in January	1618	24.0
	14.2 the second week in January	1422	21.0
	14.3 the third week in January	2106	31.2
	14.4 the fourth week in January	1824	27.0
	N. R.	1738	25.8
15. & 16. In addition to the two national meetings, there are one or more meetings each year of each Section of the Association. Please check one item in each column for each statement: I attend meetings:	*National*		
	15.1 regularly	696	10.3
	15.2 occasionally	1543	22.9
	15.3 infrequently	1820	27.0
	15.4 never	2185	32.4
	N. R.	504	7.4
	Sectional		
	16.1 regularly	1229	18.2
	16.2 occasionally	1644	24.4
	16.3 infrequently	1823	27.0
	16.4 never	1717	25.4
	N. R.	339	5.0
17. & 18. The level of mathematics presented in the programs is, for me	*National*		
	17.1 too high	1002	14.8
	17.2 about right	3162	46.9
	17.3 too low	115	1.7
	N. R.	2469	36.6
	Sectional		
	18.1 too high	525	7.8
	18.2 about right	3655	54.2
	18.3 too low	496	7.3
	N. R.	2072	30.7
19. & 20. The balance between talks or panel discussions on strictly mathematical topics and talks or panels on other topics, such as education, social implications of mathematics, mathematical applications, is	*National*		
	19.1 too heavily weighted toward mathematical topics	750	11.1
	19.2 about right	2953	43.8
	19.3 too heavily weighted toward other topics	227	3.4
	N. R.	2818	41.7
	Sectional		
	20.1 too heavily weighted toward mathematical topics	630	9.4
	20.2 about right	3206	47.5
	20.3 too heavily weighted toward other topics	373	5.5
	N. R.	2539	37.6

21.	If there were an annual or biennial survey of available text and reference books in print at the undergraduate and beginning graduate levels, would you	21.1 find such a list of little value	745	11.0
		21.2 use a library or department copy	4546	67.4
		21.3 pay \$5 (say) for a copy?	1077	16.0
		N. R.	380	5.6
22.	If there were a cumulative bibliography of selected articles on mathematics education, would you	22.1 find it of little value	1552	23.0
		22.2 use it in the library	4023	59.6
		22.3 pay \$7 (say) for a copy?	820	12.2
		N. R.	353	5.2
23.	If there were a cumulative bibliography of expository articles, would you	23.1 find it of little value	725	10.7
		23.2 use it in the library	4441	65.8
		23.3 pay \$7 (say) for a copy?	1246	18.5
		N. R.	336	5.0
24. & 25	If the MONTHLY were to publish brief abstracts of current articles from other publications			
	a. on mathematical education, would you find these	24.1 of little value	1628	24.1
		24.2 some value	3186	47.2
		24.3 great value	1693	25.1
		N. R.	241	3.6
	b. on mathematical exposition, would you find these	25.1 of little value	848	12.6
		25.2 some value	3663	54.3
		25.3 great value	1959	29.0
		N. R.	278	4.1
26.	If the Association were to publish further books of reprints of selected articles on various topics (like "Selected Papers in Calculus"), would you	26.1 find these of little value	819	12.1
		26.2 use them in the library	2900	43.0
		26.3 pay \$5 (say) for certain ones	2582	38.3
		N. R.	447	6.6
27.	If there were an encyclopedia of undergraduate mathematics, would you	27.1 find this of little value	1278	19.0
		27.2 use it in the library	4111	60.9
		27.3 pay \$30 (say) for it?	975	14.4
		N. R.	384	5.7
28.	Check the phrase which best expresses your opinion of the effect which the work of the Committee on the Undergraduate Program in Mathematics (CUPM) has had on collegiate mathematics	28.1 Strongly beneficial	1726	25.6
		28.2 Moderately beneficial	2903	43.0
		28.3 Negligible	355	5.3
		28.4 Adverse	173	2.5
		28.5 No opinion	1337	19.8
		N. R.	254	3.8

29. How has the work of CUPM affected you in your professional life?	29.1 Has helped a great deal	1082	16.0
	29.2 Has been moderately helpful	2365	35.0
	29.3 Has had little effect	1546	23.0
	29.4 Has had an adverse effect	68	1.0
	29.5 Has been largely unrelated to my professional life	1359	20.0
	N. R.	328	4.9

3. A look at certain responses. Now for some comments on individual items.

Re 2.1: One member pointed out that care should be exercised in weighing this response since "Ph. D. granting institutions also employ a great number of teachers who have nothing to do with the Ph. D. program."

One might ask if the distribution observed in question 2 matches the actual distribution of the total MAA membership. This question cannot be answered using the current membership files. I have suggested that we consider making an appropriate minor modification of the individual dues notice card in order to obtain this information from each member in the future. (This, as well as certain other demographic inquiries, will be made in time.)

Re 2.3, 6, 7, 8, 9: To me it seems we should have more members among those who are employed in the two-year colleges. Judging by the 80.4% response to #8.1, for example, and the fact that the leadership is taking immediate steps to provide the options called for, I would guess that we will attract more of our colleagues from these schools. (Until recently, there existed a legal problem in connection

TABLE A (Question 9, Expanded)

REPLY COMBINATIONS

9.1	9.2	9.3	Percentage of Responses
No reply			10.2
	X		2.7
	X	X	1.7
		X	2.1
X			41.6
X	X		19.7
X		X	5.6
X	X	X	16.4
			100.0

with the Association's taking over the TWO-YEAR COLLEGE MATHEMATICS JOURNAL. These legal problems have now been solved and, in any case, the Association indeed has plans to take over the TYCMJ no later than 1975. This may happen in 1974.)

If we look at the responses to question 9, we observe, of course, that the percentages have a total greater than 100, because the respondents checked various combinations of the MONTHLY, MATHEMATICS MAGAZINE, and the TYCMJ, respectively. Now, Table A, above, gives us the percentages for the various reply combinations received. This kind of information should be useful in the study of the journal options which are to come into being.

Re 4.1, 28, 29: The heavy response to 4.1 is borne out by the responses to 28 and 29, where we see that large numbers of the respondents have found the work of CUPM to be at least moderately helpful to them in their own professional lives. (Though NSF funding of CUPM in its present form is coming to an end, there is a widespread disposition to find ways and means to continue the work to which CUPM has been devoted.) Further comment on this is to be found in section 4, below.

Re 7.2 and 9.3: In view of the relatively low number of our members employed in two-year colleges, the response to 9.3 is quite a strong one. In view of the heavy response to 7.2 and in response to numerous inquiries received, here is the address of the TYCMJ, for those who seek information about this journal (which is edited by Professor J. Hashisaki):

The TYCMJ
53 State Street
Boston, Massachusetts 02109.

Re 10 and 11: Noteworthy changes in the MONTHLY were put into effect several years ago, after a survey conducted under the leadership of D. Bernstein. Without that survey, and the actions it led to, under Editor H. Flanders, there might have been a disheartening response to these questions.

Re 12: The decision to hold our Summer meetings two weeks before Labor Day, beginning in 1973, is consistent with the responses to 12.2.

Re 13 and 14: Here there is a two to one preference for holding our annual meetings in January. As a result of the responses to 14.3 and 14.4, we shall usually be meeting in the third week in January, and we shall occasionally meet in the fourth week. Further comment on this matter also appears in section 4, below.

Re 15, 16, 17, 18: About 33% of our members at least occasionally attend our national meetings. Though the levels of the national and sectional programs are apparently about right, I find it a bit sad that over 50% of our members attend our sectional meetings at most infrequently. (Perhaps we should have a section by section tabulation of the responses to 18.)

Re 22 and 26: Each of these two cases is, in its own way, conclusive. The MAA volume on "Selected Papers in Calculus" has been very well received and, according to 26.3, 38.3% of our members would lay out personal cash for more volumes like that one. Here, it is nice to be able to report that the Committee on Publications is authorizing a similar collection of *pre-calculus* papers.

Following the response to 22.3, for example, other projects will be given much lower priorities.

Re 27: Personally, I was a bit surprised by the magnitude of the response to 27.3. Apparently *some* of us still have money! One member did comment that encyclopaedias *do* tend to get out of date rather rapidly.

Re 28: Nearly 69% support is registered for the opinion that CUPM has had at least a moderately beneficial effect on collegiate mathematics. (See also the comment, above, Re 4.1, 28, 29.) Many grateful comments regarding CUPM came in with the completed questionnaires.

As a matter of fact, of those who *expressed an opinion* in response to question 28, nearly 90% said that the work of CUPM had been either strongly or moderately beneficial to collegiate mathematics.

4. Certain matrices. Some natural questions arise which cannot be answered by the simple tabulations in section 2, above. For example, if we look at the responses to question 6, we may well wonder how the votes are distributed among the various constituencies of our membership (as listed in question 2). We are thus led to asking our computer to give us Matrix I, below.

MATRIX I

	No Reply	Q 6.1	Q 6.2		Totals-R ows
No reply	36	25	6	.	67
Q 2.1	828	513	306	.	1647
Q 2.2	1076	1015	332	.	2423
Q 2.3	59	378	50	.	487
Q 2.4	146	166	27	.	339
Q 2.5	275	163	91	.	529
Q 2.6	149	101	46	.	296
Q 2.7	250	156	77	.	483
Q 2.8	35	28	10	.	73
Q 2.9	196	153	55	.	404
Totals-Columns	3050	2698	1000		

The matrix says, for example, that 378 respondents checked the answers 2.3 and 6.1. That is, 378 of our respondents from the two-year faculties favor the option in question, while only 50 do not. In fact, Matrix I tells us that every one of the nine constituencies listed in question 2 favors the answer 6.1 over the answer 6.2.

Now, Matrix II, below, also arises quite naturally. As indicated, this matrix is related to Matrix I, but restricts its consideration to those respondents who had indicated, by checking 7.1, that they had indeed *seen* the TYCMJ. Here too, for example, we see that the entries in the 6.1 column are greater than the corresponding elements in the 6.2 column.

MATRIX II (Restricted to Forms with Response. . .Q7.1)

	No reply	Q 6.1	Q 6.2		Totals-Rows
No Reply	0	6	1	.	7
Q 2.1	7	154	54	.	215
Q 2.2	15	271	76	.	362
Q 2.3	12	251	35	.	298
Q 2.4	1	16	0	.	17
Q 2.5	0	16	10	.	26
Q 2.6	0	12	10	.	22
Q 2.7	3	24	15	.	42
Q 2.8	0	3	0	.	3
Q 2.9	4	24	10	.	38
Totals-Columns	42	777	211	.	1030

Still referring to Matrix II, and introducing notation which the reader may easily divine, we have

$$\frac{n(7.1 \cap 6.1 \cap 2.1)}{n(7.1 \cap 2.1)} = \frac{154}{215} = 71.6\%,$$

$$\frac{n(7.1 \cap 6.1 \cap 2.2)}{n(7.1 \cap 2.2)} = \frac{271}{362} = 74.9\%,$$

$$\frac{n(7.1 \cap 6.1 \cap 2.3)}{n(7.1 \cap 2.3)} = \frac{251}{298} = 84.2\%.$$

These numbers tell us that, among those who've actually seen the TYCMJ, the sentiment in favor of taking over the journal is "uniformly" high in the university, college, and two-year college segments of our membership.

Incidentally, we also see that,

$$\frac{n(7.1 \cap 6.1)}{n(7.1)} = \frac{777}{1030} = 75.4\%.$$

Other matrices can be computed from the stored information. As various colleagues study the questions related to journal options, for example, they may well call for relevant matrices. (Should it be found expensive to compute the matrices using all 6748 responses, sampling matrices might be used.) Reproduced here, with some minimal comment, are a few more interesting matrices already in hand.

Related to the subject treated in Table A, above, are the Matrices III and IV

MATRIX III

	No reply	Q 9.1	Q 9.2	Q 9.3		Totals-rows
No reply	19	45	22	14	.	100
Q 2.1	165	1454	437	211	.	2267
Q 2.2	242	2068	1108	615	.	4033
Q 2.3	23	322	226	397	.	968
Q 2.4	38	250	218	140	.	646
Q 2.5	52	442	223	91	.	808
Q 2.6	38	239	110	57	.	444
Q 2.7	38	432	186	103	.	759
Q 2.8	5	62	47	22	.	136
Q 2.9	69	310	154	90	.	623
Totals-Columns	689	5624	2731	1740		

MATRIX IV (Restricted to Forms with Response. . .Q 8.1)

	No reply	Q 9.1	Q 9.2	Q 9.3		Totals-rows
No reply	11	37	20	12	.	80
Q 2.1	103	1076	392	198	.	1769
Q 2.2	160	1711	1004	587	.	3462
Q 2.3	18	294	213	382	.	907
Q 2.4	27	228	203	136	.	594
Q 2.5	40	355	204	84	.	683
Q 2.6	25	188	97	51	.	361
Q 2.7	27	363	169	94	.	653
Q 2.8	5	55	45	22	.	127
Q 2.9	39	250	138	83	.	510
Totals-Columns	455	4557	2485	1649		

which follow. In both of these matrices, we see, for example, that the two-year staff members record a "preference" for the TYCMJ over the MONTHLY and MATHEMATICS MAGAZINE, with MATHEMATICS MAGAZINE coming in third with them (since, respectively, $397>322>226$ and $382>294>213$).

If, for example, we now look at the 2.4 row in Matrix V below, we see (from $233>103$, primarily) that for roughly two-thirds of our high school teacher members the MONTHLY articles are at a level which is "too high." On the other hand, from the 2.2 row, we see that these articles are at a level which is "about right" for the "four-year college" members.

MATRIX V

	No reply	Q 11.1	Q 11.2	Q 11.3		Totals-rows
No reply	7	17	43	0	.	67
Q 2.1	45	150	1383	69	.	1647
Q 2.2	61	666	1662	34	.	2423
Q 2.3	13	329	145	0	.	487
Q 2.4	2	233	103	1	.	339
Q 2.5	23	159	340	7	.	529
Q 2.6	6	81	202	7	.	296
Q 2.7	14	92	368	9	.	483
Q 2.8	2	40	31	0	.	73
Q 2.9	20	124	252	8	.	404
Totals-Columns	193	1891	4529	135		

In Matrix VI, below, the elements in the 12.2 row are all maxima in their respective columns, telling us that in *all five* of the membership cases listed in question 5, the members prefer to have our summer meetings *two* weeks before Labor Day. (As noted earlier, these meetings will be scheduled that way beginning in 1973.)

MATRIX VI

	No reply	Q 5.1	Q 5.2	Q 5.3	Q 5.4	Q 5.5		Totals-rows
No reply	500	608	160	49	294	173	.	1784
Q 12.1	471	1124	177	77	582	257	.	2688
Q 12.2	642	1576	211	115	762	305	.	3611
Q 12.3	414	833	139	62	507	196	.	2151
Totals-Columns	2027	4141	687	303	2145	931		

Similar phenomena, with respect to the *Annual* meetings, occur in Matrices VII and VIII, which follow. In Matrix VII, the elements in the 13.1 row are all maxima in their respective columns. Thus a *January* meeting is preferred, with the 13.2 elements coming in last in each responsive column except for the 5.1 column (which refers to AMS members). If we look at the 5.1 column, and notice the high "no preference" count, the December choice is defeated 3 to 1 (1644 + 611 to 789). See also the comment associated with Matrix IX, below.

MATRIX VII

	No reply	Q 5.1	Q 5.2	Q 5.3	Q 5.4	Q 5.5		Totals-rows
No reply	220	211	59	18	117	71	.	696
Q 13.1	636	1644	196	108	805	313	.	3702
Q 13.2	338	789	102	52	432	151	.	1864
Q 13.3	510	611	175	57	478	196	.	2027
Totals-Columns	1704	3255	532	235	1832	731		

In Matrix VIII, it is the *third week* in January which is preferred by the members in all five listed categories, except for the *stand-off* between 14.3 and 14.4 registered by the Computing Machinery (5.2) people.

MATRIX VIII

	No reply	Q 5.1	Q 5.2	Q 5.3	Q 5.4	Q 5.5		Totals-rows
No reply	555	685	182	69	415	203	.	2109
Q 14.1	395	841	112	53	405	150	.	1956
Q 14.2	323	764	118	54	367	155	.	1781
Q 14.3	441	1151	152	79	567	246	.	2636
Q 14.4	379	963	153	72	509	219	.	2295
Totals-Columns	2093	4404	717	327	2263	973		

MATRIX IX (Restricted to Forms with Response. . .Q 5.1)

	No reply	Q 14.1	Q 14.2	Q 14.3	Q 14.4	Q 14.5		Totals-rows
No reply	188	9	9	3	8	0	.	217
Q 13.1	93	367	527	842	588	0	.	2417
Q 13.2	170	332	113	133	193	0	.	941
Q 13.3	234	133	115	173	174	0	.	829
Totals-Columns	685	841	764	1151	963	0		

The comment associated with Matrix VII, above, helps push us toward computing Matrix IX, which is a tabulation restricted to AMS members (5.1). Here the number $842 = n\ (5.1 \cap 13.1 \cap 14.3)$ is a maximum (maximorum, even), and, if we also look at the sum $2417 + 829$, from the last column, we become convinced that the third week in January is an overwhelming choice for our Annual meetings.

The last three matrices we look at are concerned with opinions on CUPM. As we see in Table B, below, about 90% of our respondents to question 28 indicate their opinion that the effect of CUPM's work on collegiate mathematics has been at least "moderately beneficial."

MATRIX X

	No reply	Q 28.1	Q 28.2	Q 28.3	Q 28.4	Q 28.5		Totals-rows
No reply	9	12	24	2	2	18	.	67
Q 2.1	58	343	806	122	66	252	.	1647
Q 2.2	45	886	1165	106	60	161	.	2423
Q 2.3	11	144	250	30	12	40	.	487
Q 2.4	19	60	117	13	2	128	.	339
Q 2.5	31	74	144	25	13	242	.	529
Q 2.6	15	47	95	7	4	128	.	296
Q 2.7	26	78	167	22	3	187	.	483
Q 2.8	8	4	19	3	2	37	.	73
Q 2.9	32	78	116	25	9	144	.	404
Totals-Columns	254	1726	2903	355	173	1337		

MATRIX XI

	No reply	Q 29.1	Q 29.2	Q 29.3	Q 29.4	Q 29.5		Totals-rows
No reply	10	6	16	15	0	20	.	67
Q 2.1	86	172	587	508	18	276	.	1647
Q 2.2	54	636	1143	450	28	112	.	2423
Q 2.3	10	123	229	96	7	22	.	487
Q 2.4	20	34	112	88	1	84	.	339
Q 2.5	21	21	53	118	4	312	.	529
Q 2.6	8	17	33	64	0	174	.	296
Q 2.7	60	27	112	109	3	172	.	483
Q 2.8	22	3	6	16	0	26	.	73
Q 2.9	37	43	74	82	7	161	.	404
Totals-Columns	328	1082	2365	1546	68	1359		

TABLE B

Percentage distribution of replies, by occupation, of those respondents to question 28 *who expressed an opinion* (i. e., those who checked 28.1, 28.2, 28.3, or 28.4).

Occupation	Number of Respondents	Strongly Beneficial	Moderately Beneficial	Negligible	Adverse
University Faculty	1337	25.7	60.3	9.2	4.8
Four-Year College Faculty	2217	40.0	52.7	4.7	2.6
Two-Year College Faculty	436	33.1	57.4	6.9	2.6
Secondary School Teachers	192	31.3	61.0	6.7	1.0
Industry	256	28.9	56.3	9.8	5.0
Government	153	30.7	62.1	4.6	2.6
Graduate Students	270	28.9	61.9	8.1	1.1
Undergraduates	28	14.3	67.9	10.7	7.1
Other	228	34.2	50.9	10.9	4.0

Of the 3990 University, four-year college and two-year college faculty who expressed an opinion on the work of CUPM,

34.5% said "strongly beneficial",
55.4% said "moderately beneficial",
6.6% said "negligible",
3.5% said "adverse".

5. A gleaning of comments to close by. We close with a few comments generated by the survey.

There was repeated favorable comment on the apparent responsive disposition of the MAA leadership. "How about establishing, on a trial basis, a 'rap session' at an Annual meeting at which members could quiz officers and/or governors about Association and/or professional matters? Or would such a session only attract cranks?"

"How about an annual (but inexpensive) publication on this matter: Just what can I do with a degree in mathematics?"

"Almost no undergraduate students are aware of such publications as PROFESSIONAL OPPORTUNITIES IN MATHEMATICS, GUIDEBOOK TO DEPARTMENTS, and other materials useful for postgraduate planning."

"Is there, could there be, a service for copying out-of-print mathematical books?"

"Could there be more textbook surveys published? With ratings of, say, elementary books?"

"Could there be a section (in the MONTHLY) on mathematical games?"

"'Educationese' should have no place in the MONTHLY."

"The readability level of the MONTHLY articles should be like that which occurs in the MAA book on selected calculus papers."

"Could the MONTHLY publish a survey article on unsolved problems?"

"How about a 'Letters to the Editor' column in the MONTHLY?"

"Remember that the MONTHLY should be for *undergraduate* math."

"Can we have more articles about the mathematicians themselves?"

(A collection of comments related to the MONTHLY has been forwarded to Editor Flanders, a selection he himself made when he viewed the correspondence.)

"Does it violate the 'freeze' to charge $5 for the Combined Membership List?"

"Does raising the dues from $10 to $12.50 violate the freeze?"

"The correct name of the organization is The Association for Computing Machinery."

"Fire ——————."

"Congratulate ——————."

"*Barbarus hic ego sum, quia non intelligor ulli.*"

"The MAA books are such good books... at low prices. But the delivery is slow."

"Can we have something like the Student Affiliates of the American Chemical Society?"

"The CUPM pamphlets were invaluable to a small, inexperienced department. I would like to see more members of CUPM with experience teaching freshmen who are not highly gifted. Most of the course outlines have too much material. Otherwise, they are admirable."

"The option of the journal would allow a department like ours (two Religious whose dues are paid by the College) to have more variety with a lower cost, less duplication. Too often only one membership can be justified unless the journals differ."

"I believe that the football player, Dick Butkus, once said, 'If I had the intelligence I would have been a brain surgeon. Since I do not, I am a middle linebacker.' For my part, if..., I would have been a Ph. D. mathematician doing research. Since ... I am not, I write computer programs and read the AMERICAN MATHEMATICAL MONTHLY."

No (further) comment.

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

A MATRIX THEORETIC CONSTRUCTION OF MAGIC SQUARES

C. R. JOHNSON, California Institute of Technology

An $n \times n$ matrix whose entries consist of the integers 1 through n^2 is called a magic square if all row and column sums are equal. There are various methods for constructing such squares; for example, the generalized uniform step method of [2], and several more or less systematic methods mentioned in [1] and [3]. This note describes a matrix theoretic method for constructing magic squares of any odd order and mentions an extension for certain even orders.

NOTATION: (a) n will denote an arbitrary odd positive integer throughout. (b) P_m will denote the group of $m \times m$ permutation matrices. (c) Q_n will denote the element of P_n with all ones on the superdiagonal (and thus a one in the lower left corner). (d) R_n will denote the symmetric $n \times n$ matrix whose first row is $0, 1, 2, \cdots, (n-1)$ and whose succeeding rows are obtained by "circulating" the first backwards. For example,

$$R_3 = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix}.$$

It is immediate to note that (1) Q_n has multiplicative order n and has no two powers $0 \leqq i < j \leqq n-1$ with a one in the same entry, and that (2) if in an arbitrary matrix $A = (a_{kl})$ the n left to right diagonals are the n vectors

$$(a_{1,1+t}, a_{2,2+t}, \cdots, a_{n-t,n}, a_{n-t+1,1}, \cdots, a_{n,t}),\ t = 0, \cdots, (n-1),$$

then each of the integers 0 through $(n-1)$ occurs on each of the n left to right diagonals of R_n.

THEOREM 1. *The matrix $M_n = [\sum_{i=0}^{n-1} (ni+1)Q_n^i] + R_n$ is an n-th order magic square.*

Proof: It suffices to show that each of the row and column sums of M_n is $n(n^2+1)/2$ and that each of 1 through n^2 occurs as an entry of M_n.

The row and column sums of $\sum_{i=0}^{n-1} (ni+1)Q_n^i$ are $\sum_{i=0}^{n-1}(ni+1)$ by note (1) above, and those of R_n are clearly $\sum_{i=0}^{n-1} i$ by construction. Thus the row and column sums of M_n are

$$\sum_{i=0}^{n-1} [(ni+1)+i] = (n+1)\sum_{i=0}^{n-1} i + \sum_{i=0}^{n-1} 1 = \frac{(n+1)(n-1)n}{2} + n = \frac{n(n^2+1)}{2}$$

as required.

Since the t-th left to right diagonal of R_n (counting from the left beginning with 0) is added to the t-th such diagonal of $\sum_{i=0}^{n-1}(ni+1)Q_n^i$ (which is just the nonzero entries of $(nt+1)Q_n^t$), the t-th diagonal of M_n runs through a complete residue system modulo n, because of note (2) above. The entries on the t-th diagonal lie between $(nt+1)$ and $n(t+1)$, so that each of 1 through n^2 occurs in M_n. This completes the proof. (By appropriately varying the weights in the sum and the definitions of Q_n and R_n additional distinct squares can be created by the same construction.) An alternate proof could be given by noting that M_n is regular in the sense of [**2**].

Using Theorem 1, one can quickly write down odd order magic squares. For example,

$$M_3 = I_3 + 4Q_3 + 7Q_3^2 + R_3$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 4 & 0 \\ 0 & 0 & 4 \\ 4 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 7 \\ 7 & 0 & 0 \\ 0 & 7 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 9 \\ 8 & 3 & 4 \\ 6 & 7 & 2 \end{bmatrix}.$$

Similarly,

$$M_5 = \begin{bmatrix} 1 & 7 & 13 & 19 & 25 \\ 22 & 3 & 9 & 15 & 16 \\ 18 & 24 & 5 & 6 & 12 \\ 14 & 20 & 21 & 2 & 8 \\ 10 & 11 & 17 & 23 & 4 \end{bmatrix} \text{ and } M_7 = \begin{bmatrix} 1 & 9 & 17 & 25 & 33 & 41 & 49 \\ 44 & 3 & 11 & 19 & 27 & 35 & 36 \\ 38 & 46 & 5 & 13 & 21 & 22 & 30 \\ 32 & 40 & 48 & 7 & 8 & 16 & 24 \\ 26 & 34 & 42 & 43 & 2 & 10 & 18 \\ 20 & 28 & 29 & 37 & 45 & 4 & 12 \\ 14 & 15 & 23 & 31 & 39 & 47 & 6 \end{bmatrix}.$$

Since arbitrary interchanges of rows or columns cannot affect row and column sums, we have the following theorem concerning any magic square and the permutation group of its order.

THEOREM 2. *If M is any $m \times m$ magic square and if $P \in P_m$, then PM and MP are m-th order magic squares.*

Of course, there are several even more trivial methods of obtaining "different" magic squares from M, such as transposing and rotating. The essential difference is that these methods cannot change the relative positions of the entries of M.

The square M_n is not magic in its two main diagonals (i.e., they do not have the same sums as the rows and columns). However, it would be of interest if there were such a square among the orbit of M_n under P_n (i.e., $\{PM_n\colon P \in P_n\}$). Indeed, this is the case, and one such square can in general be found explicitly.

Define T_n to be the blockwise direct sum of those matrices in $P_{(n+1)/2}$ and $P_{(n-1)/2}$ which have all ones along the antidiagonal, and take the larger of the two to be in the upper left hand corner of T_n. For instance

$$T_5 = \left[\begin{array}{ccc|cc} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

Now $T_n \in P_n$ and we have the following:

THEOREM 3. *The magic square T_nM_n is also magic in its two main diagonals.*

Proof: An inspection of M_n shows that the n numbers

$$\frac{n+1}{2}, \frac{3n+1}{2}, \frac{5n+1}{2}, \cdots, \frac{(2n-1)n+1}{2}$$

occur on the $(n+1)/2$ st right to left diagonal and the n numbers

$$\frac{n(n-1)}{2} + 1, \frac{n(n-1)}{2} + 2, \cdots, \frac{n(n+1)}{2}$$

occur on the $(n+1)/2$ st left to right diagonal. The sum of each set of numbers is the required $n(n^2+1)/2$ and a simple computation shows that under T_n they are transformed to the two main diagonals. As an illustration,

$$T_3M_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 & 9 \\ 8 & 3 & 4 \\ 6 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 3 & 4 \\ 1 & 5 & 9 \\ 6 & 7 & 2 \end{bmatrix}.$$

Theorems 1 and 2 provide an easy method for constructing a class of $(n!)^2$ magic squares of order n, and Theorem 3 exhibits a distinguished element within this class.

With a slight additional effort one may inductively construct squares of even orders which are not powers of two from these given squares. The method is reminiscent of classical ones and places an appropriately ordered 2×2 block consisting of $4(i-1)+1$ through $4(i-1)+4$ in the position where i occurs in the previously constructed square.

References

1. W. S. Andrews, Magic Squares and Cubes, 2nd edition, Open Court Publishing, Chicago, 1917,

2. T. M. Apostol, and H. S. Zuckerman, On magic squares constructed by the uniform step method, Proc. Amer. Math. Soc., 2 (1951) 557–565.

3. B. Rosser, and R. J. Walker, The algebraic theory of diabolic magic squares, Duke Math. J. 5 (1939) 705–728.

GROUPS WHOSE ELEMENTS ARE OF ORDER TWO OR THREE

E. D. BOLKER, Bryn Mawr College

In this note we characterize those groups all of whose elements are of order two or three and which contain at least one element of each kind. Call such a group **acceptable.** There are two classes of acceptable groups: some resemble S_3, the symmetric group on three symbols, the others A_4, the alternating group on four. The result, which I state precisely below, is not new; it was first proved by B. H. Neumann in [**1**] and used by him to settle the Burnside conjecture for $k = 3$: every finitely generated group all of whose elements have order $\leqq k = 3$ is finite. I rediscovered Neumann's theorem while solving a special case of a problem posed in this MONTHLY [**2**]: Characterize those pairs $A < G$ ("$<$" means "is a subgroup of") for which for all x, $A \cup \{x,x^{-1}\} < G$. When $A = \{e\}$, $A \cup \{x,x^{-1}\} < G$ just when x has order two or three. To solve the problem then means to characterize acceptable groups. There are two reasons for publishing this new proof. First, it is easy and elementary. The little the reader needs to know about group extensions is explained in the course of the argument. Second, recent progress has been made on characterizing groups whose elements have orders less than or equal to five, so it seemed worthwhile to have this easier case accessible.

Let G be a group. Write S (T) for the set of elements of G of order two (three) and, when $R \subseteq G$, write R^* for $R \cup \{e\}$. Then G is acceptable when neither S nor T is empty and $G = S^* \cup T$. Before we can characterize acceptable groups, we must study two almost acceptable cases.

Suppose T is empty, so that every element of G has order two. Then G is abelian and is naturally a vector space over the field $\mathbb{Z}_2$, so that it is characterized by its dimension d. Let Γ be a set of cardinality d; then G is isomorphic to $\perp\!\!\!\perp_\Gamma \mathbb{Z}_2$, the group of functions from Γ to $\mathbb{Z}_2$ each of which is 0 except at finitely many points of Γ.

If S is empty, so that all elements are of order three, then G is said to have exponent three. Finding all such groups is nontrivial. If, however, G is abelian, then it is easy to verify that it is naturally a vector space over $\mathbb{Z}_3$ and hence is just $\perp\!\!\!\perp_\Gamma \mathbb{Z}_3$; the cardinality of Γ determines G. We shall need to know later that, whether or not G is abelian, if it has more than three elements then it contains a subgroup isomorphic to $\mathbb{Z}_3 \times \mathbb{Z}_3$.

We prove that it suffices to find a nontrivial pair of commuting elements. If we knew that G had a finite subgroup with more than three elements that would follow from the well-known nontriviality of the center of such a group. But without that knowledge we proceed as follows. Since G has more than three elements, we can find $x, y \neq e$ with $x \neq y, y^{-1}$. If x and y do not commute, then we shall show that xy and yx do. First note that, by assumption, $e \neq xy \neq yx$. Moreover,

$xy \neq (yx)^{-1}$ because $xy = (yx)^{-1} = x^{-1}y^{-1} = x^2y^2$ implies $e = xy$. Finally, xy and yx commute because

$$(xyyx)(yxxy)^{-1} = (xy^2x)(y^2xy^2) = (xy^2)^3 = e.$$

Now we can build all the acceptable groups.

Groups of type T. Let Γ be a set of given cardinality and let $H = \perp\!\!\perp_\Gamma \mathbb{Z}_3$. The map sending each element of H to its inverse is an automorphism of order two, so we can form the semidirect product (splitting extension) $G = H \circledS \mathbb{Z}_2$ determined by this automorphism: G is the set $H \times \{\pm 1\}$ with multiplication $\langle h,a\rangle\langle k,b\rangle = \langle hk^a, ab\rangle$. Then it is easy to see that G is an acceptable group in which $T^* = H < G$. When Γ has one element, G is isomorphic to S_3.

Groups of type S. Let Γ be a set of given cardinality, V the Klein four-group and $K = \perp\!\!\perp_\Gamma V$. A cyclic permutation α of the three nonidentity elements of V is an automorphism of order three of V and hence determines such an automorphism of K. Let G be the semidirect product $K \circledS \mathbb{Z}_3$ determined by this action. That is, G is the set $K \times \mathbb{Z}_3$ with multiplication $\langle h,a\rangle\langle k,b\rangle = \langle h \cdot \alpha^a(k), a+b\rangle$, where we think of $\mathbb{Z}_3$ as $\{0,1,2\}$ under addition modulo three. Then G is an acceptable group in which $S^* = K < G$. When Γ has one element, G is isomorphic to A_4.

We shall show that every acceptable group is of type S or T. We write $a, b, c, \cdots$ (resp. $\cdots x, y, z$) for elements of S (resp. T). When p and q commute, write $p \sim q$. Our argument begins with some elementary observations, clearly true in groups of types S or T, which we prove for an arbitrary acceptable group.

1. $a \nsim x$. (If $ax = xa$, then ax has order six, a contradiction.)
2. $a \sim b \Leftrightarrow ab \in S^*$. ($ab = ba \Rightarrow (ab)^2 = a^2b^2 = e \Rightarrow ab \in S^* \Rightarrow ab = (ab)^{-1} = b^{-1}a^{-1} = ba$.)

Note that in groups of type S we always have $a \sim b$, while in groups of type T, $a \neq b$ implies $a \nsim b$. This motivates the next observation.

3. $\sim$ is transitive on S. (If $ab = ba$ and $bc = cb$ then $b \sim ac$. Hence $ac \notin T$ (#1) so $ac \in S^*$ and thus $a \sim c$ (#2).)
4. $a \nsim b \Rightarrow ab \in T$ (#2) $\Rightarrow ababab = e \Rightarrow aba = bab$.
5. $ay \in S \Rightarrow ayay = e \Rightarrow aya^{-1} = aya = y^{-1}$.
6. $x \sim y \Rightarrow (xy)^3 = x^3y^3 = e \Rightarrow xy \in T^*$.

LEMMA. *If G is acceptable, then either $S^* < G$ or $T^* < G$.*

Proof. If every pair of elements of S commutes then $S^* < G$, and conversely (#2), so suppose there is a noncommuting pair and $S^* \nless G$. We shall show that no two distinct elements of S commute. For $a \in S$, let C_a be the centralizer of a. Then $C_a \subset S^*$ (#1) and $S^* \neq C_a < G$, for if they were equal, S^* would be a subgroup of G. Suppose $b \nsim a$ and $c \sim a$; we shall show $c = a$ or e. If $c \neq e$, then

since $\sim$ is transitive, $c \nsim b$. Let $d = bcab$. Then

$$(*) \qquad \begin{aligned} da &= bc(aba) \\ &= (bcb)ab \qquad (\#4) \\ &= cbcab \qquad (\#4 \text{ again}) \\ &= cd. \end{aligned}$$

Now $d^2 = e$ because $a \sim c$. If $d \sim a$ then (*) implies $a = c$, while if $d \nsim a$ then

$$\begin{aligned} da &= adad \quad (\#4) \\ &= acdd \quad (*) \\ &= ac = ca \end{aligned}$$

so $d = c \sim a$, a contradiction.

Now we can show T^* is closed under multiplication. If $xy \notin T^*$, then $xy \in S$ so $xyxy = e$ and $yxyx = y(xyxy)y^{-1} = e$ and $yx \in S$ as well. We must have $x \nsim y$ lest $xy \in T^*$ (#6) so $xy \neq yx$ and hence $xy \nsim yx$. Then

$$z = xyyx = xy^2x \notin S^* \ (\#2) \text{ so } e = z^3 = xy^2x^2(y^2x^2)y^2x.$$

But $y^2x^2 = y^{-1}x^{-1} = (xy)^{-1} = xy$ so substituting in the last equation yields

$$e = z^3 = xy^2x^2(xy)y^2x = z$$

so $z = e$, a contradiction. Thus $xy \in T^*$ and $T^* < G$.

THEOREM. *Every acceptable group is of type S or T.*

Proof. We shall show that if S^* $(T^*) < G$ then G is of type S (T). Suppose $T^* < G$. If $a \in S$ and $y \in T$ then $ay \notin T^*$ lest a be in T^*, which is closed under multiplication. Thus $ay \in S$, so, fixing $a \in S$ and applying #5, we see that the map $y \rightsquigarrow aya^{-1} = y^{-1}$ is an automorphism of T^*. Hence T^* is abelian and so is a product $\perp\!\!\perp_\Gamma \mathbb{Z}_3$. Now suppose $a, b \notin T^*$. Then $abyb^{-1}a^{-1} = ay^{-1}a^{-1} = y$ so $ab \sim y$. Thus $ab \in T^*$, so T^* is of index two in G, which is therefore a semidirect product of T^* with $\mathbb{Z}_2$, with the induced action $y \rightsquigarrow y^{-1}$ making G of type T.

Suppose, on the other hand, that $S^* < G$. Since S^* is abelian, it is a product $\perp\!\!\perp_\Delta \mathbb{Z}_2$. S^* is normal in G; let K be any subgroup of G/S^*. Then K acts on S^* by conjugation. Let R be an orbit of that action; we shall show $R^* < S^*$. Suppose $a, b \in R \neq \{e\}$ and $a \neq b$. Then there is a $y \in G$ with $yS^* \in K$ and $yay^{-1} = b$. Let $c = yby^{-1} \in R$; then $a = ycy^{-1}$ since y has order three. Then $y(abc)y^{-1} = bca = abc \in S^*$. Since $y \sim abc$, #1 implies $abc = e$, so $ab = c$. Thus $R^* < S^*$. Moreover, since no y fixes an $a \in R$, $\#R = \#K$. Now if G/S^* had more than three elements, we could take for K a nine element subgroup and thus produce a

ten element subgroup of S^*. Since every such subgroup has order a power of two, we must have G/S^* isomorphic to $\mathbb{Z}_3$ and for each orbit $R \neq \{e\}$ of the action of $\mathbb{Z}_3$ on S^*, R^* is isomorphic to V and $R^* \circledS \mathbb{Z}_3$ is isomorphic to A_4 and hence is of type S.

Call a family $\{R_\gamma\}_{\gamma \in \Gamma}$ of orbits **independent** if in the subgroup H of S^* they generate, each element has a unique expansion $\prod_{\gamma \in \Gamma} a_\gamma$ where $a_\gamma \in R_\gamma^*$ and $a_\gamma = e$ for almost all γ. Then $H \circledS \mathbb{Z}_3$ is of type S. Let Γ index a maximal independent family. Then H is invariant under the action of $\mathbb{Z}_3$ on S^*. If it were a proper subgroup of S^* there would be an orbit R disjoint from H and $\{R_\gamma\} \cup \{R\}$ would be a larger independent family. Thus $H = S^*$ and G is of type S.

References

1. B. H. Neumann, Groups whose elements have bounded orders, J. London Math. Soc., 12 (1937) 195.

2. A. P. Street, Advanced Problem # 5742, this MONTHLY, 77 (1970) 655, and 78 (1971) 799.

SUMS OF FINITE SETS OF INTEGERS

MELVIN B. NATHANSON, Southern Illinois University

Let $\mathscr{A}$ be a finite set of integers. The h-fold sum of $\mathscr{A}$, denoted by $h\mathscr{A}$, is the set of all sums of h elements of $\mathscr{A}$, repetitions being allowed. In this note we describe exactly all sufficiently high sums of any finite set of integers.

All latin letters stand for integers, and script letters for finite sets of integers. Denote by $(a_1, a_2, \cdots, a_k)$ the greatest common divisor of $a_1, a_2, \cdots, a_k$. Let $[p, q]$ be the set of integers n such that $p \leqq n \leqq q$. Let $z - \mathscr{D} = \{z - d \mid d \in \mathscr{D}\}$ and $z + \mathscr{D} = \{z + d \mid d \in \mathscr{D}\}$.

THEOREM. *Let $\mathscr{A} = \{a_0, a_1, \cdots, a_k\}$ be a finite set of integers with $a_0 = 0 < a_1 < \cdots < a_k = a$ and $(a_1, a_2, \cdots, a_k) = 1$. Then there exist non-negative integers C and D and sets $\mathscr{C} \subset [0, C-2]$ and $\mathscr{D} \subset [0, D-2]$ such that for all $h \geqq a^2 k$*

$$h\mathscr{A} = \mathscr{C} \cup [C, ha - D] \cup ha - \mathscr{D}. \tag{1}$$

We require the following lemma:

LEMMA. *Let $a_1, a_2, \cdots, a_k = a$ be positive integers with $(a_1, a_2, \cdots, a_k) = 1$. Assume that*

$$(a-1) \sum_{i=1}^{k-1} a_i \leqq n \leqq ha - (k-1)(a-1)a.$$

Then there exist non-negative integers $u_1, u_2, \cdots, u_k$ such that

$$n = u_1 a_1 + u_2 a_2 + \cdots + u_k a_k$$

and

$$\sum_{i=1}^{k} u_i \leqq h.$$

Proof. Since $(a_1, a_2, \cdots, a_k) = 1$, there are integers $x_1, x_2, \cdots, x_k$ such that

$$n = x_1 a_1 + x_2 a_2 + \cdots + x_k a_k.$$

For $i = 1, 2, \cdots, k-1$, let u_i be the least non-negative residue of x_i modulo a_k. Then

$$\begin{aligned} n &\equiv x_1 a_1 + x_2 a_2 + \cdots + x_{k-1} a_{k-1} \pmod{a_k} \\ &\equiv u_1 a_1 + u_2 a_2 + \cdots + u_{k-1} a_{k-1} \pmod{a_k} \end{aligned}$$

and so there exists an integer u_k such that

$$n = u_1 a_1 + u_2 a_2 + \cdots + u_{k-1} a_{k-1} + u_k a_k.$$

From the lower and upper bounds assumed on n, it follows easily that $u_k \geqq 0$ and $\sum_{i=1}^{k} u_i \leqq h$.

Proof of the Theorem. Let $H = a^2 k$. Let $[C, Ha - D]$ be the largest interval of integers such that

$$\left[(a-1) \sum_{i=1}^{k-1} a_i,\ Ha - (k-1)(a-1)a\right] \subset [C, Ha - D] \subset H\mathscr{A}.$$

The lemma asserts that such an interval exists. Let $\mathscr{C} = H\mathscr{A} \cap [0, C-2]$ and $Ha - \mathscr{D} = H\mathscr{A} \cap [Ha - D + 2, Ha]$. Then $\mathscr{D} \subset [0, D-2]$ and $H\mathscr{A} = \mathscr{C} \cup [C, Ha - D] \cup Ha - \mathscr{D}$. Thus (1) holds for H.

The theorem is proved by induction on h. Suppose that (1) is true for some $h \geqq H$. Let

$$\begin{aligned} \mathscr{B} &= \mathscr{C} \cup [C, (h+1)a - D] \cup (h+1)a - \mathscr{D} \\ &= \mathscr{C} \cup [C, C + a - 1] \cup [C + a, (h+1)a - D] \cup (h+1)a - \mathscr{D}. \end{aligned}$$

By the lemma, $C + D \leqq Ha \leqq ha$, and so the second equality holds. We must show that $(h+1)\mathscr{A} = \mathscr{B}$.

Observe that

$$C \leqq (a-1) \sum_{i=1}^{k-1} a_i < a^2 k = H \leqq h \tag{2}$$

and

$$ha - D - C \geqq Ha - (k-1)(a-1)a - (a-1) \sum_{i=1}^{k-1} a_i \geqq a. \tag{3}$$

Since $0 \in \mathscr{A}$ and $a \in \mathscr{A}$, it follows that $h\mathscr{A} \subset (h+1)\mathscr{A}$ and $a + h\mathscr{A} \subset (h+1)\mathscr{A}$.

Then $\mathscr{C} \subset h\mathscr{A} \subset (h+1)\mathscr{A}$. Inequality (3) implies that

$$[C, C+a-1] \subset [C, ha-D] \subset h\mathscr{A} \subset (h+1)\mathscr{A}.$$

Similarly,

$$[C+a, (h+1)a-D] = a + [C, ha-D] \subset (h+1)\mathscr{A}$$

and $(h+1)a - \mathscr{D} = a + (ha - \mathscr{D}) \subset (h+1)\mathscr{A}$. Therefore, $\mathscr{B} \subset (h+1)\mathscr{A}$.

Let $b \in (h+1)\mathscr{A}$. If $b < C$, then inequality (2) implies that b cannot be the sum of $h+1$ nonzero elements of $\mathscr{A}$, so $b \in h\mathscr{A}$, hence $b \in \mathscr{C} \subset \mathscr{B}$. If $C \leqq b < C+a$, then $b \in [C, C+a-1] \subset \mathscr{B}$.

Suppose that $b \in (h+1)\mathscr{A}$ and $b \geqq C+a$. It suffices to show that $b - a \in h\mathscr{A}$. Then, by the induction hypothesis (1), either

$$b \in a + [C, ha-D] = [C+a, (h+1)a-D] \subset \mathscr{B}$$

or

$$b \in a + (ha - \mathscr{D}) = (h+1)a - \mathscr{D} \subset \mathscr{B},$$

and so $(h+1)\mathscr{A} \subset \mathscr{B}$, hence $(h+1)\mathscr{A} = \mathscr{B}$. But if $b - a \notin h\mathscr{A}$, then b is the sum of $h+1$ elements of $\mathscr{A}$ which are all less than a. Thus,

$$b \leqq (h+1)(a-1). \tag{4}$$

But by (1) we have $[C, ha-D] \subset h\mathscr{A}$, and so the conditions $b - a \geqq C$ and $b - a \notin h\mathscr{A}$ imply that

$$b - a > ha - D \geqq ha - (k-1)(a-1)a. \tag{5}$$

Inequalities (4) and (5) give $h < (k-1)(a-1)a - 1 < a^2k = H$, which is absurd. Therefore, $(h+1)\mathscr{A} \subset \mathscr{B}$, and the proof of the theorem is complete.

It is an easy exercise to show that $C = 0$ if and only if $a_1 = 1$, and that $D = 0$ if and only if $a_{k-1} = a_k - 1$.

Clearly, an arbitrary finite set of integers differs from the normalized sets considered in the theorem only by a translation and contraction.

This research was supported in part by an NSF Predoctoral Traineeship from the University of Rochester.

A WEAK PARALLELOGRAM LAW FOR l_p

W. L. BYNUM AND J. H. DREW, College of William and Mary

A vector space V with norm $\| \quad \|$ *obeys a weak parallelogram law* (or *is a weak parallelogram space*, or briefly, *is a w. p. space*) if there is a γ, $0 < \gamma \leqq 1$, such that for all x, $y \in V$,

$$\|x+y\|^2 + \gamma\|x-y\|^2 \leqq 2\|x\|^2 + 2\|y\|^2 \tag{1}$$

(see [4]). The interest in this inequality stems from the well-known theorem of Jordan and von Neumann, which says that a Banach space satisfies (1) for $\gamma = 1$ if and only if it is a Hilbert space (see [2, p. 115] for a discussion of the Jordan-von Neumann theorem and subsequent results).

Let l_p denote the set of all infinite sequences $x = \{x_k\}$ of real numbers such that $\Sigma |x_k|^p < \infty$, with the norm $\|x\| = (\Sigma |x_k|^p)^{1/p}$. In [4] it was shown that for each n, the subspace $l_p(n)$ is a w. p. space, where $l_p(n)$ is the set of all $x = \{x_k\}$ in l_p such that $x_k = 0$ for $k > n$. However, the methods used in [4] were not adequate to show that l_p itself is a w. p. space. In fact, no infinite dimensional w. p. spaces were known. This paper establishes the following theorem:

THEOREM. *If $1 < p \leqq 2$, then l_p is a w. p. space; moreover, the largest possible value of γ in* (1) *for l_p is $p - 1$.*

The proof is based on the following theorem of Hanner. We give an outline of the proof for reference.

THEOREM (Hanner [3]). *If $1 < p \leqq 2$ and x and y are in l_p, then*

$$(\|x\| + \|y\|)^p + \big|\|x\| - \|y\|\big|^p \leqq \|x + y\|^p + \|x - y\|^p. \tag{2}$$

Proof. First, we show that it suffices to prove (2) for non-negative sequences in l_p. This entails showing (as you will see later) that if u and v are complex numbers, then

$$(|u| + |v|)^p + \big||u| - |v|\big|^p \leqq |u + v|^p + |u - v|^p. \tag{3}$$

The right side of (3) can be rewritten as: $(a^2 + b^2 + 2abt)^{p/2} + (a^2 + b^2 - 2abt)^{p/2}$, where $a = |u|$, $b = |v|$, and $-1 \leqq t \leqq 1$. This expression, as a function of t, has a minimum on the interval $[-1, 1]$ at 1 and -1; moreover, this minimum is the left side of (3).

Now, let x and y be in l_p and let x^* and y^* be sequences such that for each n, $x_n^* = |x_n|$ and $y_n^* = |y_n|$. Then, $\|x\| = \|x^*\|$, and $\|y\| = \|y^*\|$, and by (3), $\|x^* + y^*\|^p + \|x^* - y^*\|^p \leqq \|x + y\|^p + \|x - y\|^p$. Thus, it suffices to show that (2) holds for x^* and y^*.

To do this, set $q = 1/p$ and introduce the function g as follows:

$$g(u,v) = (u^q + v^q)^p + |u^q - v^q|^p \qquad (u, v \geqq 0).$$

Note that for $t \geqq 0$, $g(tu, tv) = tg(u,v)$. For each n, let $a_n = (x_n^*)^p$ and $b_n = (y_n^*)^p$. Then, we can rewrite (2) as follows:

$$g(\Sigma a_n, \Sigma b_n) \leqq \Sigma g(a_n, b_n). \tag{4}$$

To obtain (4), we need only establish that

$$g(a + b, c + d) \leqq g(a,c) + g(b,d) \qquad (a,b,c,d \geqq 0). \tag{5}$$

To establish (5), consider the function $h(t) = g(t, 1)$ for $t \geqq 0$. Since h' is everywhere continuous and increasing, h is convex. If $c > 0$, and $d > 0$, then $(c + d)\, h((a + b)/(c + d)) \leqq c\, h(a/c) + d\, h(b/d)$, which is precisely (5). If c or d is 0, a similar argument establishes (5).

The following lemma is also needed in the proof of our theorem.

LEMMA. *If $1 < 1 + \gamma \leqq p \leqq 2$ and if a and b are real numbers, then*

$$(a + b)^2 + \gamma(a - b)^2 \leqq 2^{2-(2/p)}(|a|^p + |b|^p)^{2/p}.$$

Proof. Let $k = 2 - (2/p)$. For t real, let

$$h(t) = 2^k(1 + |t|^p)^{2/p} - (1 + t)^2 - \gamma(1 - t)^2.$$

Since $h(t) = t^2 h(1/t)$ for $t \neq 0$ and since $h(t) \geqq h(|t|)$, it is sufficient to show that $h \geqq 0$ on $[0, 1]$. For x in $(0, \frac{1}{2}]$, let $g(x) = (2 - p)x^k + (p - 1)\, x^{k-1}$. Then, $g'(x) \leqq 0$ and therefore, g is decreasing on $(0, \frac{1}{2}]$. For t in $(0, 1]$,

$$h''(t) = 2^{k+1}\, g(t^p/(1 + t^p)) - 2 - 2\gamma,$$

so h'' is decreasing on $(0, 1]$. Since $h''(1) = 2(p - 1 - \gamma) \geqq 0$, h' is increasing on $[0, 1]$. But $h'(1) = 0$. Thus, on $[0, 1]$, h is decreasing and $h \geqq h(1) = 0$.

Proof of the theorem. Suppose x and y are in l_p. By letting $2a = \| x + y \| + \| x - y \|$ and $2b = \| x + y \| - \| x - y \|$, we can rewrite the inequality of Hanner's Theorem as: $a^p + |b|^p \leqq \| x \|^p + \| y \|^p$. Let $\gamma = p - 1$ and $k = 2 - (2/p)$. By the lemma, Hanner's theorem, and the Hölder inequality:

$$\begin{aligned}(a + b)^2 + \gamma(a - b)^2 &\leqq 2^k(a^p + |b|^p)^{2/p} \leqq 2^k(\| x \|^p + \| y \|^p)^{2/p} \\ &\leqq 2^k 2^{1-k}(\| x \|^2 + \| y \|^2).\end{aligned}$$

For $t \geqq 0$, let x_t be the member of l_p whose first two coordinates are $1 + t$ and $1 - t$, with remaining coordinates zero. Let $y = x_0$. Using l'Hôpital's theorem twice, we obtain that

$$(2\| x_t \|^2 + 2\| y \|^2 - \| x_t + y \|^2)/\| x_t - y \|^2 \to p - 1$$

as $t \to 0$. Thus, $p - 1$ is the largest possible value of γ in (1) for the space l_p.

As Lindenstrauss has noted in [**5**, p. 243], the proof of Hanner's theorem in [**3**] is valid for a general measure space $L_p(\mu)$ (the set of functions f on a set X, measurable with respect to a σ-ring of subsets of X, such that $|f|^p$ is μ-integrable). Moreover, if there are at least two disjoint subsets of X of positive finite measure, then we can show as above that $p - 1$ is the largest possible w. p. constant for $L_p(\mu)$.

The theorem of this paper also answers in the negative the following two questions posed at a conference on functional analysis held at Sopot, Poland, in 1968:

(i) Is each w. p. space with an unconditional Schauder basis isomorphic (linearly homeomorphic) to l_2?

(ii) Is each closed subspace Y of a w. p. space X complemented in X? That is, does there exist a closed subspace Z of X such that for each x in X there is a unique y in Y and a unique z in Z such that $x = y + z$?

A counterexample for both (i) and (ii) is $l_p(1 < p < 2)$. Question (i) has a negative answer because l_p and l_2 are of incomparable linear dimension; that is, neither is isomorphic to a closed subspace of the other [**1**, Theorem 7, p. 205]. The answer to (ii) is negative because l_p has an uncomplemented closed subspace [**6**, p. 138].

References

1. S. Banach, Théorie des Opérations Linéaires, 2nd. ed., Chelsea, New York, 1963 (1933).

2. M. M. Day, Normed Linear Spaces, Springer-Verlag, Berlin, 1962.

3. O. Hanner, On the uniform convexity of L_p and l_p, Ark. Mat., 3(1956) 239–244.

4. D. C. Kay, A parallelogram law for certain L_p spaces, this MONTHLY, 74(1967) 140–147.

5. J. Lindenstrauss, On the modulus of smoothness and divergent series in Banach spaces, Mich. Math. J., 10(1963) 241–252.

6. F. J. Murray, On complementary manifolds and projections in spaces L_p and l_p, Trans. Amer. Math. Soc., 41 (1937) 138–152.

A LOWER BOUND FOR AN AREA INTEGRAL

D. J. NEWMAN, Yeshiva University

In a recent issue of this publication, [1], C. K. Chui asked whether there existed $c > 0$ such that, for any $z_1, z_2, \cdots, z_n$ on the unit circle, we have

$$(1) \qquad \iint\limits_{|z|<1} \left| \sum_{\nu=1}^{n} \frac{1}{z - z_\nu} \right| dA \geqq c \quad (dA \text{ the area measure}).$$

He even suggested the possibility that this integral is minimized by the choice $z_\nu = e^{2\pi i\nu/n}$ so that we would have

$$(2) \qquad \iint\limits_{|z|<1} \left| \sum_{\nu=1}^{n} \frac{1}{z - z_\nu} \right| dA \geqq \iint\limits_{|z|<1} \left| \frac{nz^{n-1}}{1 - z^n} \right| dA$$

and (2) easily implies (1).

Although we are unable to give a proof for the attractive conjecture, (2), we find that we can indeed prove (1), and that is the purpose of this note.

Proof of (1). Let $P_\nu = \text{Re}(z_\nu + z)/(z_\nu - z)$, $\nu = 1, 2, \cdots, n$, call S_ν the set where $P_\nu \geqq 2n$ and write χ_ν as the characteristic function of S_ν. Finally call $S = \bigcup_{\nu=1}^{n} S_\nu$.

Since

$$(3) \qquad \sum_{\nu=1}^{n} \frac{1}{z - z_\nu} = -\frac{1}{2z} \left(\sum_{\nu=1}^{n} \frac{z_\nu + z}{z_\nu - z} - n \right)$$

we certainly have

$$\left|\sum_{\nu=1}^{n} \frac{1}{z - z_\nu}\right| \geqq \frac{1}{2}\left(\sum_{\nu=1}^{n} P_\nu - n\right) \tag{4}$$

so that

$$\iint_{|z|<1} \left|\sum_{\nu=1}^{n} \frac{1}{z - z_\nu}\right| dA \geqq \frac{1}{2} \iint_{S} \left(\sum_{\nu=1}^{n} P_\nu - n\right) dA. \tag{5}$$

Next observe that since $P_\nu \geqq 0$ we certainly have $P_\nu \geqq P_\nu \chi_\nu$ and that throughout S, $\Sigma \chi_\nu \geqq 1$. Hence, throughout S, we have $\Sigma_{\nu=1}^{n} P_\nu - n \geqq \Sigma_{\nu=1}^{n} (P_\nu - n)\, \chi_\nu$. Finally we may observe that $(P_\nu - n)\, \chi_\nu \geqq n\, \chi_\nu$ (since $P_\nu \geqq 2\, n$ unless $\chi_\nu = 0$) and this allows us to write, throughout S,

$$\sum_{\nu=1}^{n} P_\nu - n \geqq n \sum_{\nu=1}^{n} \chi_\nu.$$

Inserting this into our integral gives the lower bound

$$\iint_{|z|<1} \left|\sum_{\nu=1}^{n} \frac{1}{z - z_\nu}\right| dA \geqq \frac{n}{2} \iint \sum_{\nu=1}^{n} \chi_\nu\, dA. \tag{6}$$

But each of the sets S_ν is a disc of radius $1/(2n+1)$ included inside (and tangent to) the unit circle. Hence

$$\iint \chi_\nu\, dA = \pi/(2n+1)^2$$

and so

$$\frac{n}{2} \iint \cdot \sum_{\nu=1}^{n} \chi_\nu\, dA = \frac{\pi n^2}{2(2n+1)^2} \geqq \frac{\pi}{18}. \tag{8}$$

Thus (1) is proved with $c = \pi/18$.

Reference

1. C. K. Chui, A lower bound of fields, this MONTHLY, 78 (1971) 779–780.

BAIRE FUNCTIONS AND EXTREME POINTS

LAWRENCE G. BROWN, Stanford University

$C(X)$, the space of continuous complex-valued functions on the compact Hausdorff space X, is a well-known Banach space. If

$$U = \{f \in C(X) : |f(x)| \leqq 1,\ \forall x \in X\}$$

and

$$E = \{f \in C(X) : |f(x)| = 1,\ \forall x \in X\},$$

then U is the unit ball of $C(X)$, and E is the set of extreme points of U. At least three papers ([2], [3], [4]) deal with the theorem that U is the closed convex hull of E. None of these papers uses what we consider to be the simplest proof of the theorem, and perhaps the reason is that the measure-theoretic lemma involved is not as well known as it should be. We present this lemma and show how it can be used to prove the theorem. We are grateful to G. M. Leibowitz for advice on this subject.

LEMMA 1. *If X is a compact (or locally compact, σ-compact) Hausdorff space and M an arcwise-connected separable metric space, then the set of Baire functions from X to M is $\mathscr{S}$, the smallest set of functions containing the continuous functions and closed under sequential pointwise convergence.*

Proof. The definition of Baire function, of course, is that $f^{-1}(A)$ be a Baire set [1] in X for each Borel set A in M. The lemma is standard for the special case $M = [0,1]$ (see [1], p. 223, Exercise 6), and the general case follows from this.

We first prove it is sufficient to show that $\mathscr{S}$ contains the simple functions. Let f be a Baire function, $\varepsilon > 0$, and $M = \bigcup_{n=1}^{\infty} M_n$ a decomposition of M into disjoint Borel sets of diameter less than ε. Choose $p_n \in M_n$, and define f_ε by: $f_\varepsilon(x) = p_n$ if $f(x) \in M_n$. Since $f_\varepsilon(x) \to f(x)$ as $\varepsilon \to 0$, it is clearly sufficient to show $f_\varepsilon \in \mathscr{S}$. Now define g_N by:

$$g_N(x) = \begin{cases} p_n & \text{if } f(x) \in M_n \text{ for } n \leqq N \\ p_1 & \text{otherwise.} \end{cases}$$

Clearly, g_N is simple, and $g_N(x) \to f_\varepsilon(x)$.

Now if f is simple, there is a continuous function $\phi: [0,1] \to M$ such that $\phi([0,1]) \supset f(X)$. Thus there is a Baire function $f_0: X \to [0,1]$ such that $\phi \circ f_0 = f$. Let $\mathscr{S}_0 = \{g_0: g_0: X \to [0,1] \text{ and } \phi \circ g_0 \in \mathscr{S}\}$. Clearly, $\mathscr{S}_0$ contains all continuous functions, and is closed under sequential pointwise convergence. By the special case $M = [0,1]$, $f_0 \in \mathscr{S}_0$, and hence $f \in \mathscr{S}_0$.

REMARKS. 1. If we are given a Baire measure on X, then it follows that every Baire function is the pointwise almost everywhere limit of a sequence of continuous functions.

2. It may be of interest to find necessary and sufficient conditions on the separable metric space M for the conclusion of Lemma 1 to hold for arbitrary X. It is necessary but not sufficient that M be connected. It is sufficient but not necessary that M have a dense arcwise connected component. If M is a topological group, the latter is also necessary. If M is a locally compact group, connectedness implies that the arc-component of the identity be dense; and hence connectedness is necessary and sufficient. In general, a necessary and sufficient condition is that, for every finite $F \subset M$ and $\varepsilon > 0$, there is an arc-component which comes within ε of each point in F; but this condition appears unwieldy.

3. Those X for which the conclusion holds for arbitrary (still separable metric) M are precisely the totally disconnected ones.

4. If M is a non-separable metric space, then any continuous function from X to M has separable range. Under the continuum hypothesis, the same would be true of Baire functions, and then the separability would not be needed.

LEMMA 2. *If μ is a finite complex Baire measure on the compact Hausdorff space X, then* $|\mu|(X) = \sup\{|\int f d\mu| : f \in E\}$.

Proof. Let $S = \sup\{|\int f d\mu| : f \in E\}$. Let $\nu = |\mu|$, and write $d\mu = \rho d\nu$ where ρ is a Baire function and $|\rho(x)| \equiv 1$. Let

$$\mathscr{S}' = \{f : f \text{ is a Baire function, } |f(x)| \equiv 1, \text{ and } |\int f d\mu| \leqq S\}.$$

By Lemma 1, for the case where M is the circle, $\bar{\rho} \in \mathscr{S}'$. Hence $|\mu|(X) = \nu(X) = \int \bar{\rho} d\mu \leqq S$.

The theorem now follows from Lemma 2, the Riesz Representation Theorem, and the double polar theorem (see any textbook on linear topological spaces).

References

1. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, N. J., 1950.

2. N. T. Peck, Representation of functions in $C(X)$ by means of extreme points, Proc. Amer. Math. Soc., 18 (1967) 133–135.

3. R. R. Phelps, Extreme points in function algebras, Duke Math. J., 32 (1965) 267–277.

4. R. Sine, On a paper of Phelps, Proc. Amer. Math. Soc., 18 (1967) 484–486.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

AN EDGE-COLOURING PROBLEM

NORMAN BIGGS, Royal Holloway College, London.

1. The footballers of Croam. In the little English hamlet of Croam the consuming passion of the inhabitants is Association Football. In fact, the members of the village football team have become so ruthless in their will to win that no other team will play against them.

Thus the eleven footballers of Croam (who are, incidentally, the only able-bodied men in the village) are forced to arrange their own matches between two teams of five, with the eleventh man as referee. Further, such is the bitterness of recrimination which follows even these matches, that it has proved necessary to rule that only one match can be played with the same teams and the same referee. This rule was originally regarded with some misgiving, as it was felt that it might seriously limit the number of matches which could be played. However, a villager who has a head for figures worked out that there are 1386 different ways of splitting the eleven men into two teams of five and a referee. This number is thought to be adequate but not generous, for the footballers of Croam are dedicated men.

But there is a second rule which these men, united by their love of football, but embittered by isolation, have been forced to make in order to keep the peace. No five men will play together as a team more than once on any given day of the week. Therein lies the problem. Can all the possible matches be played under this restriction? Can all the matches be played if Sunday games are not allowed?

2. Commentary. The problem as stated is the case $k = 6$ of the following general situation. Let X denote a set of odd finite cardinality $2k - 1$ and let V denote the set of subsets of X having exactly $k - 1$ members. Construct the graph O_k whose vertex set is V and in which two vertices are joined by an edge if and only if they are disjoint subsets of X; it follows that O_k is a regular graph of valency k. How many colours are needed to colour the edges of O_k in such a way that adjacent edges have different colours?

It is clear that at least k colours are necessary, and by the powerful general result due to Vizing **[2]**, $k + 1$ colours are sufficient for any graph of valency k. Thus Vizing's theorem gives an immediate answer to the first part of our problem of the footballers: the graph O_6 of valency 6 whose vertices represent teams of five men and whose edges represent matches can be edge-coloured with seven colours, and so all the matches can be played using the seven days of the week. The crucial part of our problem, however, is the second question, which becomes: Can O_6 be edge-coloured with 6 colours?

In general it seems hard to find regular graphs of valency k which cannot be edge-coloured with k colours, unless we insist that the number of vertices be odd. For if an edge-k-colouring exists, then the set of edges of any particular colour covers each vertex precisely once and, since each edge is incident with two vertices, the total number of vertices must be even. This remark disposes of our general problem when

$$|V| = \binom{2k-1}{k-1}$$

is odd, and this is so if and only if k is a power of 2. (The reader who prefers numbers to graphs may digress to prove this statement, and its generalisation that the expo-

nent of 2 in the binomial coefficient is the number of ones in the binary expansion of k, less one.)

The case $k = 3$ corresponds to a graph with ten vertices; this graph is Petersen's graph [1] which is one of the few known examples of a trivalent graph which is not edge-colourable with three colours. Thus O_k has no edge-k-colouring when k is 3 or a power of 2, and we conjecture that this is so for all k.

References

1. J. Petersen, Die Theorie der regulären Graphen, Acta Math., Stockholm, 15 (1891) 193–220.
2. V. G. Vizing, On an estimate of the chromatic class of a p-graph, (Russian) Diskret. Analiz., 3 (1964) 25–30.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

PICARD'S THEOREM

JAMES FABREY, University of North Carolina

Picard's Theorem on the existence and uniqueness of solutions of ordinary differential equations is frequently stated without proof in elementary courses. First-order equations are often treated by standard methods [1], but a proof for *arbitrary* order is postponed until systems of equations have been studied. This is not necessary if one considers only *linear* equations [2]. This is not a severe restriction in an elementary course, and first-order techniques easily generalize. A uniqueness proof has already been outlined [3, p. 12–13]. This article provides an existence proof, together with an approximation method that is a worthwhile teaching device (even if the proof is bypassed).

THEOREM. *Suppose f and a_i, $1 \leqq i \leqq n$, are continuous functions on an open interval J which contains the origin, and b_k, $0 \leqq k < n$, are constants. Then*

$$Ly = D^n y + \sum_{i=1}^{n} a_i D^{n-i} y = f \tag{1}$$

has a unique solution y on J which satisfies the initial data

$$D^k y(0) = b_k, \qquad 0 \leqq k < n. \tag{2}$$

REMARK. In (1) we have assumed that the equation has already been divided by the leading coefficient. In (2) we have translated the initial time to the origin; this is no loss of generality.

If $n = 1$, then the standard method of Picard is to convert

$$Ly = Dy + a_1 y = f \tag{3}$$

into an integral equation by integrating (3) once:

$$y(t) - b_0 = \int_0^t (f(s) - a_1(s)y(s))ds. \tag{4}$$

One might use Newton's method of approximate roots to motivate the recursion formula for approximate solutions:

$$y_1(t) \equiv 0, \quad y_{m+1}(t) = b_0 + \int_0^t (f(s) - a_1(s)y_m(s))ds, \quad m \geqq 1. \tag{5}$$

One then shows that y_m converges uniformly on closed sub-intervals of J to a unique solution y.

The general case is treated similarly. Let D^{-1} denote the integral operator, $\int_0^t$, and $D^{-n} \equiv (D^{-1})^n$ denote n successive integrations. For example, $D^{-2}\cos t = D^{-1}\sin t = -\cos t + 1$. Since D^n is not invertible ($D^n c = 0$ if c is a constant), this notation is somewhat misleading. However, if S is the set of all n-times differentiable functions on J whose initial data vanish, then D^n restricted to S is invertible with inverse D^{-n}. This follows from repeated applications of the Fundamental Theorem of the Calculus, and it is a good exercise for the reader.

We call $p_n(t) = \sum_{k=0}^{n-1} b_k t^k / k!$ the *initial polynomial*. It satisfies the initial data (2). By linearity, $y - p_n$ is in S whenever y satisfies (2). Moreover, $D^n p_n = 0$ so that (1) may be rewritten

$$D^n(y - p_n) = f - L'y, \tag{6}$$

where $L'y = \sum_{i=1}^{n} a_i D^{n-i} y$. Motivated by (4), we integrate (6) n times (i.e., apply D^{-n}):

$$y - p_n = D^{-n}(f - L'y). \tag{7}$$

Finally, we generalize (5):

$$y_1 = 0, \quad y_{m+1} = p_n + D^{-n}(f - L'y_m), \quad m \geqq 1. \tag{8}$$

Example. Let $Ly = D^2 y - y, f = 0, y(0) = Dy(0) = 1$. Then it is easy to compute $y_m(t) = \sum_{j=0}^{2m-3} t^j / j!$, $m \geqq 2$, so that $e^t = \lim_{m\to\infty} y_m(t)$ is the desired solution.

The reader should be warned that, just as with the first-order case, the approximate solutions and limits may frequently be difficult to compute. We conclude with an existence proof that is completely standard except for (9). We shall prove that y_m

converges uniformly to a solution on every closed sub-interval $[a,b]$ of J which contains the origin. Since every point in J is contained in such a set, we may solve (1) and (2) uniquely in J. The proof might be supplemented with theorems on uniform convergence.

Proof of Theorem. For $j \geqq 2$, let $r_j = y_{j+1} - y_j$. By *telescoping series*,

$$D^k y_m = D^k y_2 + \sum_{j=2}^{m-1} D^k r_j, \quad 0 \leqq k \leqq n, \quad m \geqq 3.$$

Suppose that there exist positive constants α and K such that for t in $[a,b]$,

$$|D^k r_j(t)| \leqq \alpha K^{j-2}/(j-2)!. \tag{9}$$

Then $\sum_{j=2}^{\infty} |D^k r_j(t)| \leqq \alpha e^K$, so the limits $\lim_{m\to\infty} y_m = y$ and $\lim_{m\to\infty} D^k y_m$ exist and the convergence is uniform on $[a,b]$. Therefore, y is n-times differentiable, and $D^k y = \lim_{m\to\infty} D^k y_m$, $0 \leqq k \leqq n$. Since each y_m satisfies (2), so does y. By uniform convergence, we may interchange limits with D^{-n} and L' to obtain (6). Thus, y is a solution with correct initial data.

It remains to establish (9). By (8), $r_j = -D^{-n}L' r_{j-1}$. Hence, by the Fundamental Theorem of the Calculus,

$$D^k r_j = -D^{k-n} L' r_{j-1}, \ D^n r_j = -L' r_{j-1}, \qquad 0 \leqq k < n. \tag{10}$$

Thus, it suffices to bound

$$B_j(t) = \max_{0 \leqq k < n} |D^k r_j(t)|, \ C_j(t) = |L' r_j(t)|$$

by $\alpha K^{j-2}/(j-2)!$ on $[a,b]$. By (10), we obtain the estimates

$$B_j(t) \leqq \max_{0 \leqq k < n} |D^{k-n} C_{j-1}(t)|, \ C_j(t) \leqq nMB_j(t), \tag{11}$$

where M is the maximum absolute value of the coefficients of L' on $[a,b]$. Let c be the maximum of B_2 on $[a,b]$, and let d be the maximum of $|t|, |t|^2, \cdots, |t|^{n-1}$ on $[a,b]$. We iterate (11) to obtain $C_2(t) \leqq cnM$ and

$$B_3(t) \leqq cnM \max_{0 \leqq k < n} \frac{|t|^{n-k}}{(n-k)!} \leqq cnMd|t|, \ C_3 \leqq c(nM)^2 d|t|,$$

$$B_4(t) \leqq c(nM)^2 d \max_{0 \leqq k < n} \frac{|t|^{1+n-k}}{(1+n-k)!} \leqq c(nM)^2 d^2 \frac{|t|^2}{2!}, \ C_4 \leqq c(nM)^3 d^2 \frac{|t|^2}{2!}.$$

It is easy to show by induction that

$$B_j(t) \leqq c(nMd)^{j-2} \frac{|t|^{j-2}}{(j-2)!}, \ C_j(t) \leqq c(nM)^{j-1} d^{j-2} \frac{|t|^{j-2}}{(j-2)!}.$$

The proof is completed with $K = nMd(b-a)$, $\alpha = \max\{c, cnM\}$.

References

1. F. Brauer and J. Nohel, Ordinary Differential Equations, Benjamin, New York, 1967.
2. J. Fabrey, Linear Differential Equations, Allyn and Bacon, Boston, to appear.
3. A. Rabenstein, Introduction to Ordinary Differential Equations, Academic Press, New York, 1966.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

MATHEMATICS FOR THE CAPTURED STUDENT

S. K. STEIN, University of California, Davis

A few months ago Professor Dilworth asked me if I would be interested in speaking about mathematics for the student who takes it as a requirement, not for cultural reasons. I said I would, and began to visit other colleges and to survey the recent literature on the problem. I focused my attention on two-year colleges, in part because the problem is acknowledged there frankly and clearly, and in part because four-year colleges and universities contain, whether they admit it or not, two-year colleges. Moreover, the variety of two-year colleges is almost as great as that of four-year colleges. In one two-year college 70% of the students plan on going on for a bachelor's degree, though the chief school counselor told me that much of this percentage is "fantasy planning." In another, many students are on relief, a sizeable number are ex-cons and achieving a 2-year degree may put them back in the job market in an economy that has little room for the unskilled. As one teacher put it, "The open door of the community college brings in people of a much wider range of abilities and inabilities than it did 10 or 15 years ago." And another, "There are great numbers of students showing up for arithmetic and algebra who have either forgotten everything or failed to understand anything, or who simply never had anything to do with mathematics before." And the open admission policy of such universities as C. C. N. Y. certainly will magnify the problem of the captured student there.

The "captured student" in the title refers to a wide variety of students who, much to their surprise and disappointment, are suddenly forced to study a subject they may have been fleeing for years, even if fresh out of high school. Such a student might be majoring in psychology and have to take statistics, or in home economics and have to take "some mathematics," or he may have to take mathematics for a

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before February 28, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2379. *Proposed by H. Kestelman, University College, London, England*

Find all matrices A such that both A and A^{-1} have all elements real and nonnegative.

E 2380. *Proposed by Erwin Just, Bronx Community College*

Let $f(x)$ be an irreducible polynomial of degree at least three with rational coefficients, and suppose that $f(x)$ has precisely two non-real zeros, $z_1 = p + qi$ and $z_2 = p - qi$, where p and q are real. Could q possibly be rational?

E 2381. *Proposed by E. S. Langford, University of Maine*

Suppose that $\{f_n\}$ is a sequence of continuous real-valued functions defined on $[0,1]$ such that $f_1(x) \geqq f_2(x) \geqq \cdots \geqq 0$ for all $x \in [0,1]$. Suppose further that the only continuous function f such that $f_n(x) \geqq f(x) \geqq 0$ for all $x \in [0,1]$ and all $n = 1, 2, \cdots$ is the zero function. Is it necessarily true that

$$\int_0^1 f_n(x)dx \to 0 \quad \text{as} \quad n \to \infty?$$

E 2382. *Proposed by Thomas Hughes, Fort Worth, Texas*

One has a number of balls, identical in appearance; one of the balls is known to be slightly heavy, another slightly light by the same amount, and the rest have a standard weight. It is desired to isolate both the light and heavy balls, using only three weighings on a "triple platform balance." (A triple platform balance consists of three arms forming a Y, equally spaced at intervals of 120°; these are supported at the center, and at the end of each arm is a pan. If n balls are placed in each of the three pans, then one can tell whether each of the three sets of balls is heavier, lighter, or the same weight as n standard balls; note however that the heavy ball and the light ball together weigh as much as two standard balls.)

What is the largest number of balls from which one can identify both the heavy ball and the light ball in only three weighings?

E 2383. *Proposed by E. T. Ordman, University of Kentucky*

Let n be a nonnegative integer. For $p = 1, 2, \cdots$, define

$$S_p(n) = \sum_{k=0}^{[\frac{n}{2}]} \left[\binom{n}{k} - \binom{n}{k-1}\right]^p,$$

where we make the usual conventions regarding binomial coefficients. It is easy to evaluate $S_1(n)$. Evaluate $S_2(n)$.

E 2384.* *Proposed by H. W. Gould, West Virginia University*

Using the notation of E 2383 above, show that $S_3(n)$ is always divisible by $S_1(n)$.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Difficult Triangle Inequality

E 2245 [1970, 652; 1971, 793]. *Proposed by A. W. Walker, Toronto, Canada*

If A, B, C; a, b, c; s are the angles, side lengths, and semi-perimeter of any plane triangle, then

$$(a+b+c)^3(s-a)(s-b)(s-c) \geqq (a^2+b^2+c^2)^3 \cos A \cos B \cos C.$$

II. *Comment by A. van Tooren, Leusden, Holland.* We show that the inequality

$$(*) \qquad (abc)^2(a+b+c)^3(a+b-c)(a-b+c)(-a+b+c)$$
$$\geqq (a^2+b^2+c^2)^3(a^2+b^2-c^2)(a^2-b^2+c^2)(-a^2+b^2+c^2),$$

which the proposer's solution [1971, 795] indicates is equivalent to the proposed inequality, does indeed hold for all nonnegative a, b, c which do not form the sides of a triangle. Multiplying both sides of the identity

$$(a+b+c)(-a+b+c)(a-b+c)(a+b-c) = -\sum a^4 + 2\sum a^2b^2$$

by $\sum a^2$, we obtain

$$\begin{aligned}(a^2+b^2+c^2)&(a+b+c)(-a+b+c)(a-b+c)(a+b-c)\\ &= -\sum a^6 + \sum a^4b^2 + 6a^2b^2c^2\\ &= (-a^2+b^2+c^2)(a^2-b^2+c^2)(a^2+b^2-c^2) + 8(abc)^2.\end{aligned}$$

This enables us to write the desired inequality (*) in the form

$$(a+b+c)(-a+b+c)(a-b+c)(a+b-c)\left[(abc)^2(a+b+c)^2-(a^2+b^2+c^2)^4\right] + 8(abc)^2(a^2+b^2+c^2)^3 \geqq 0.$$

We are assuming that a, b, c are nonnegative and do not form a triangle. Therefore

$$(a+b+c)(-a+b+c)(a-b+c)(a+b-c) \leqq 0.$$

To complete the proof we show that

$$(abc)^2(a+b+c)^2 - (a^2+b^2+c^2)^4 \leqq 0.$$

The case $a = b = c = 0$ is trivial. In all other cases, since the left side is homogeneous, we are allowed to assume that $a^2+b^2+c^2 = 1$. Then

$$abc(a+b+c) = a^2bc + b^2ca + c^2ab \leqq a^2+b^2+c^2 = (a^2+b^2+c^2)^2.$$

This completes the proof.

Editor's comment. Note that van Tooren's proof does not apply when a, b, c, are the sides of a triangle. A proof of (*) was received also from Dorothee Aeppli for the case covered by van Tooren. She also submitted a rather complicated direct algebraic proof of the "triangle" case. A straightforward algebraic proof of (*), valid for all nonnegative a, b, c, is still solicited.

An Inequality for the Complex Logarithm

E 2319 [1971, 1019]. *Proposed by Thomas Hern, Bowling Green State University*

If z_1 and z_2 are complex numbers with $0 < |z_1| \leqq 1$ and $0 < |z_2| \leqq 1$, show that $|z_1 - z_2| \leqq |\log z_1 - \log z_2|$.

I. *Solution by Henrik Meyer, Birkerød, Denmark.* Write $z_j = r_j\, e^{i\theta_j}$, for $j = 1,2$. By the Mean Value Theorem,

$$|\log r_1 - \log r_2| = \frac{1}{\xi}|r_1 - r_2| \geqq |r_1 - r_2|,$$

where $0 < \xi < 1$ By the Law of Cosines

$$\begin{aligned}|z_1 - z_2|^2 &= r_1^2 + r_2^2 - 2r_1r_2\cos(\theta_1 - \theta_2)\\ &= (r_1 - r_2)^2 + 2r_1r_2[1 - \cos(\theta_1 - \theta_2)]\\ &= (r_1 - r_2)^2 + 4r_1r_2\sin^2\tfrac{1}{2}(\theta_1 - \theta_2)\\ &\leqq (r_1 - r_2)^2 + r_1r_2(\theta_1 - \theta_2)^2 \leqq (r_1 - r_2)^2 + (\theta_1 - \theta_2)^2\\ &\leqq (\log r_1 - \log r_2)^2 + (\theta_1 - \theta_2)^2 = |\log z_1 - \log z_2|^2.\end{aligned}$$

(The inequality $4\sin^2\frac{1}{2}(\theta_1 - \theta_2) \leqq (\theta_1 - \theta_2)^2$ follows from $|\sin\theta| \leqq |\theta|$.)

II. *Solution by O. P. Lossers, Technological University, Eindhoven, Netherlands, and (independently) by J. B. Conway, Indiana University.* The inequality is equivalent to

$$|e^{w_1} - e^{w_2}| \leqq |w_1 - w_2|,$$

where $\operatorname{Re} w_1 \leqq 0$ and $\operatorname{Re} w_2 \leqq 0$. If γ is the straight line segment from w_1 to w_2, then

$$\begin{aligned}|e^{w_1} - e^{w_2}| = \left|\int_\gamma e^w dw\right| \leqq \int_\gamma |e^w|\,|dw|\\ \leqq |w_1 - w_2| \max_{w\in\gamma} |e^w| \leqq |w_1 - w_2|.\end{aligned}$$

III. *Generalization by R. J. Evans, Jackson State College.* We prove the stronger inequality

$$|\log z_1 - \log z_2| \geqq \frac{2|z_1 - z_2|}{|z_1| + |z_2|}$$

which holds for all nonzero z_1 and z_2. Putting $w = z_1/z_2$ reduces the inequality to

$$|\log w| \geqq \frac{2|w-1|}{1 + |w|}.$$

Write $w = re^{i\theta}$. Whenever w satisfies the above inequality, so do $\bar{w}$ and $1/w$; hence, we may assume that $r \geqq 1$ and $\theta \geqq 0$. We must show that

$$f(r,\theta) = ((\log r)^2 + \theta^2)(1 + r^2) - 4(r^2 + 1 - 2r\cos\theta) \geqq 0.$$

When $\theta \geqq 0$,

$$\frac{\partial f}{\partial \theta} = 2\theta(1 + r)^2 - 8r\sin\theta \geqq 2\theta(r - 1)^2 \geqq 0.$$

Thus $f(r,\theta) \geqq f(r,0)$. It remains to show that $f(r,0) \geqq 0$. Hence we must show that

$$g(r) = (1 + r)\log r - 2(r - 1) \geqq 0$$

for $r \geqq 1$. We have $g'(r) = \log r + (1 - r)/r$. By applying the Mean Value Theorem to $\log x$ on $[1, r]$, we see that $g'(r) \geqq 0$ for each $r \geqq 1$. Hence $g(r) \geqq g(1) = 0$.

Also solved by the proposer and forty-one other readers.

Editor's comment. Even though the complex logarithm is a multivalent function, the inequality holds in the sense that no matter which values are taken for $\log z_1$ and $\log z_2$, the inequality is valid. For fixed z_1 and z_2, it should be clear that $|\log z_1 - \log z_2|$ is smallest when "we take z_1 and z_2 as close together as possible on the Riemann surface for the logarithm;" that is, if $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$, then $|\theta_1 - \theta_2| \leqq \pi$.

Divisibility of the Numerator of a Sum of Fractions

E 2320 [1971, 1019]. *Proposed by Erwin Just, Bronx Community College*

Let

$$\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}, \cdots, \frac{a_n}{b_n}\right\}$$

consist of n rational numbers in which the a_i and b_i are integers, and $(n, \Pi_{i=1}^n b_i) = 1$. Prove that there exist positive integers k and m such that the numerator of the fraction determined by $\sum_{i=k}^m a_i/b_i$ is divisible by n.

I. *Solution by Neal Felsinger, U. S. Army.* Let $b = \Pi_{i=1}^n b_i$ and let $c_i = a_i b/b_i$. Then

$$\sum_{i=k}^{m} \frac{a_i}{b_i} = \sum_{i=k}^{m} \frac{c_i}{b}.$$

Since n and b are relatively prime, it follows that n divides $\sum_{i=k}^m c_i$ if and only if it divides the numerator of $\sum_{i=k}^m a_i/b_i$. We note that this latter fraction need not be in lowest terms as long as no additional factors (i.e., other than divisors of b) are introduced in the denominator. Thus we can assume that all denominators are equal.

Consider the n sums $s_k = \sum_{i=1}^k a_i$ for $k = 1, 2, \cdots, n$. If all are distinct modulo n, then one is a multiple of n and we are done. Otherwise, there are q and m with $q < m$ such that $\sum_{i=1}^m a_i$ and $\sum_{i=1}^q a_i$ are congruent mod n. Letting $k = q + 1$ we have that n divides $\sum_{i=k}^m a_i$.

II. *Comment by Andrzej Mąkowski, Warsaw, Poland.* Let G be the additive group of rational numbers p/q in lowest terms, where q and n are relatively prime, and let H be its subgroup consisting of all numbers with $n | p$. The problem is then a special case of Problem 4300 [1950, 47] for the group G/H.

Also solved by the proposer and twenty-five other readers.

A Summation Known to Euler

E 2321 [1971, 1020]. *Proposed by Michael Skalsky, Southern Illinois University*

Show that

$$\sum_{n=1}^{\infty} (nxe^{-x})^n/n! = x(1-x)^{-1}.$$

I. *Solution by R. G. Buschman, University of Wyoming.* If the Maclaurin series for e^{-nx} is substituted into the given series and the double series is rearranged by setting $m = k + n$ for a new summation index [i.e., by summing "diagonally" — Ed.], we have

$$\sum_{m=1}^{\infty} S_m x^m/m!, \text{ where } S_m = \sum_{k=0}^{m} (-1)^k \binom{m}{k} k^m.$$

We need only show that $S_m = m!$ to obtain a geometric series which has the desired sum. Sums of this type can be evaluated by methods given in Chapter 5, Section 12 of I. J. Schwatt, *An Introduction to the Operations with Series*, Chelsea, New York, 1961, where the series

$$S(m, p) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} k^p$$

is evaluated. We need only $S_m \equiv S(m, m) = m!$.

II. *Solution by M. L. Glasser, Battelle Memorial Institute.* Let $u = xe^{-x}$. Then by Lagrange's Theorem (Whittaker and Watson, *A Course of Modern Analysis*, Cambridge Univ. Press, 1958, p. 132) we have

$$x = \sum_{n=1}^{\infty} \frac{n^{n-1}u^n}{n!}. \tag{1}$$

Thus

$$\sum_{n=1}^{\infty} \frac{(nxe^{-x})^n}{n!} = \sum_{n=1}^{\infty} \frac{n^n u^n}{n!} = u \frac{d}{du} \left\{ \sum_{n=1}^{\infty} \frac{n^{n-1}u^n}{n!} \right\}$$

$$= u \frac{dx}{du} = u \left(\frac{du}{dx} \right)^{-1} = x(1-x)^{-1}.$$

Also solved by forty-nine other readers.

Editor's comments. Glasser's equation (1) above can be found in G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, I, Berlin, 1964, Problem 209 (p. 125); see the solution to this problem on p. 301 where (1) is credited to Euler.

Harry Pollard remarks that the formula is the special case $\alpha = 0$ of

$$\sum_{n=0}^{\infty} \frac{(n+\alpha)^n}{n!} (xe^{-x})^n = \frac{e^{\alpha x}}{1-x}, \tag{2}$$

which has an old history. See Pólya-Szegö, *op. cit.*, Exercise 214 (p. 126) and its solution (p. 302). Eldon Hansen also cites this formula and notes that it can be found in John Riordan, *Combinatorial Identities*, New York, 1968, p. 147 and in Leonard Carlitz, *The coefficients in an asymptotic expansion*, Proc. AMS 16 (1965), 248–252. He remarks that the formula is used in Problems 4776 [1958, 783] and 4868 [1960, 704].

Arnold Scheinman comments that the given problem is a special case of Jensen's formula (equivalent to (2) above)

$$\sum_{j=0}^{\infty} \frac{(a+jx)^j}{j!} e^{-(a+jx)} = \frac{1}{1-x},$$

which is a common expansion in queuing theory. He cites J. L. W. V. Jensen, ***Sur une identité d'Abel et sur autres formules analogues***, Acta Math., 26 (1902).

David Shelupsky solves the problem using Cauchy's Theorem on a suitable integral in the complex plane, and Robert Shafer and S. J. Bernau (independently) investigate the problem where x is complex.

Otto Ruehr submits a solution in which he uses Jabotinski's Theorem (E. Jabotinski, ***Representation of functions by matrices***, Proc. AMS 4 (1953)) to obtain a series expansion for x^m in terms of y, where $y = xe^{-x}$. [The case $m = 1$ is just Lagrange's formula as given in the solution by Glasser. — Ed.]

Glasser notes that the problem is given in L. Onsager and N. Samaras, Jour. Chem. Physics 2 (1934), p. 528.

Several readers solve the problem by computing the Maclauren series for $F(x) = \sum_{n=1}^{\infty} (nxe^{-x})^n/n!$ using Leibniz's Theorem, and then comparing the coefficients with those in the Maclaurin series for $x/(1-x)$.

Most "formal" series manipulations guarantee equality only for sufficiently small x. Several solvers investigate the range of x for which the given formula holds. By the ratio test or the root test using Stirling's Formula, the series must converge if $|xe^{-x}| < e^{-1}$ and diverge if $|xe^{-x}| > e^{-1}$; that is, the series must converge if $x \in (x_0, 1) \cup (1, \infty)$ and diverge if $x < x_0$, where x_0 is the solution to the transcendental equation $xe^{-x} = -e^{-1}$. [Numerical methods give $x_0 = -0.2784645428\cdots$. —Ed.] If $x = 1$, the series diverges since its terms are of the order of $n^{-1/2}$ (this can be shown by Stirling's Formula), whereas if $x = x_0$, the series converges by the alternating series test. Thus the series converges if and only if $x \in [x_0, 1) \cup (1, \infty)$. But the series cannot equal $x/(1-x)$ if $x > 1$ since the sum of the series is positive and $x/(1-x)$ is negative. The Maclaurin series for $x/(1-x)$ converges if $|x| < 1$ so that it follows that equality holds if and only if $x_0 \leqq x < 1$.

Your editors were indeed gratified by the wide variety of solutions and comments which they received for this problem. We regret that because of space limitations we are able to print only a small selection.

Unions of Finite Sets of Subsets

E 2322 [1971, 1020]. *Proposed by Harry Lass and Peter Gottlieb, Jet Propulsion Laboratory, California Institute of Technology*

Let $A_1, \cdots, A_n$ be finite sets, each with the same number of elements, and let $S = \bigcup_{j=1}^{n} A_j$. Suppose that for some fixed k with $1 \leqq k \leqq n$, every union of k of the sets is S and every union of fewer than k of the sets is a proper subset of S. Determine in terms of n and k (1) the minimum number of elements in S; (2) the number of elements in each A_j when the number of elements in S is minimal; and (3) the number of elements common to any j of the subsets when S is minimal.

Solution by D. M. Bloom, Brooklyn College. Let $N = \{1, 2, \cdots, n\}$ and for each $x \in S$ let $M(x) = \{i \colon i \in N \text{ and } x \in A_i\}$. We first show that any $(n+1-k)$-element subset T of N equals $M(x)$ for some $x \in S$. Indeed, since $|N \backslash T| = k - 1$ it follows that $\bigcup_{i \notin T} A_i$ is not all of S. Let $x \in S$ be an element which is not in this union. For each $j \in T$ we have $S = (\bigcup_{i \notin T} A_i) \cup A_j$ which implies that $x \in A_j$ and consequently $M(x) = T$.

It follows from the result above that $|S| \geqq \binom{n}{k-1}$ with equality if and only if the sets $M(x)$ are precisely the $(n+1-k)$-element subsets of N and the correspondence $x \to M(x)$ is injective. In this case

$$|A_i| = |\{x: i \in M(x)\}| = |\{M(x): i \in M(x)\}| = \binom{n-1}{n-k},$$

which is the number of $(n+1-k)$-element subsets of N containing i. Thus

$$|A_i| = \binom{n-1}{n-k} = \binom{n-1}{k-1}.$$

Conversely, to show that the case $|S| = \binom{n}{k-1}$ actually occurs, for each i we define A_i to be the set of all $(n+1-k)$-element subsets of N which contain i. Clearly $S = \bigcup_{i=1}^{n} A_i$ and $|S| = \binom{n}{k-1}$. All the hypotheses of the problem are satisfied and hence (1) and (2) are answered.

Finally if $|S|$ is minimal, then an element x belongs to a given intersection of j of the A_i if and only if the corresponding indices i all belong to $M(x)$. Hence the answer to question (3) is $\binom{n-j}{k-1}$, the number of $(n+1-k)$-element subsets of N which contain j given elements.

Also solved by Virginia Bolton, Robert Breusch, Ralph Freese, M. G. Greening (Australia), Robert Patenaude, G. S. Sidhu, D. P. Sumner, Dorothy Wolfe, and the proposers.

An Inequality for All Triangles

E 2323 [1971, 1020]. *Proposed by Anders Bager, Hjørring, Denmark*

Characterize those triangles for which $\sqrt{3} + 5\,\Sigma\cot A \geqq 3\,\Sigma\csc A$. (Here $\Sigma f(A)$ is taken to mean $f(A) + f(B) + f(C)$.)

I. *Solution by F. Leuenberger, Feldmeilen, Switzerland.* The inequality holds for any triangle with equality if and only if the triangle is equilateral. To prove this, note that $2\cot A = \cot\frac{1}{2}A - \tan\frac{1}{2}A$ and $2\csc A = \cot\frac{1}{2}A + \tan\frac{1}{2}A$ so that the statement can be written as

$$\sqrt{3} + \Sigma\cot\tfrac{1}{2}A \geqq 4\,\Sigma\tan\tfrac{1}{2}A$$

which is equivalent to

$$\sqrt{3} + \frac{s}{r} \geqq \frac{4(4R+r)}{s},$$

where s, R, r denote the semiperimeter, circumradius and inradius respectively. But since $F = rs$, this is equivalent to

$$F\sqrt{3} + s^2 \geqq 4r(4R+r),$$

where F denotes the area. It is known that $F\sqrt{3} \geqq 9r^2$ and $s^2 \geqq 16Rr - 5r^2$, with equality in each case if and only if the triangle is equilateral. The first inequality follows from item 4.2 or 7.9 of O. Bottema et al., *Geometric Inequalities*, Wolters-Noordhoff, Groningen, 1969, and the second is item 5.8, *op. cit.* Adding these two inequalities gives the desired result.

II. *Solution by Leonard Goldstone, Watervliet, N. Y.* We note that $\Sigma \cot A = \Sigma a^2/4F$ and that $\Sigma \csc A = \Sigma ab/2F$. Letting $Q = \Sigma(a-b)^2$, the given inequality is transformed into

$$4F\sqrt{3} + 3Q \geqq \Sigma a^2,$$

which is known to hold for all triangles, with equality if and only if the triangle is equilateral. (Item 4.7 of *Geometric Inequalities.*)

Also solved by Robert Breusch, Ralph Garfield, M. G. Greening (Australia), and Carolyn MacDonald.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before February 28, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.
An asterisk (*) means neither the proposer nor the editors supplied a solution.

5878. *Proposed by Václav Konečný, Jarvis Christian College, Hawkins, Texas*

Show that

$$\int_0^\infty e^{-x}\ln^2 x\,dx = \gamma^2 + \frac{\pi^2}{6} \quad\text{and}\quad K - 5/2 < \int_0^\infty e^{-x}\ln^3 x\,dx < K - 9/4,$$

where $K = -\gamma(\gamma^2 + \pi^2/2)$ and γ is Euler's constant.

5879.* *Proposed by Alexandru Lupaş, Institutul de calcul, Cluj, Romania, and the University Stuttgart, Germany*

Let $L_n: C(K) \to C(K)$, $n = 1,2,\cdots, K = [0,1]$, be a sequence of linear positive operators with the properties:

(1) $L_n e_0 = e_0$, $L_n e_1 = e_1$, $L_n e_2 = e_2 + a_n$, $(n = 1,2,\cdots)$, where $e_k(t) = t^k$ and $\{a_n\}$, $n = 1,2,\cdots$, is a sequence of nonnegative continuous functions, uniformly convergent to zero on K, and such that there is x_0, $x_0 \in K$, for which $a_n(x_0) > 0$.

(2) For every $g \in C(K)$ and $n = 1,2,\cdots$,

$$(L_n g)(0) = g(0),\ (L_n g)(1) = g(1).$$

Prove or disprove the following assertion: A function f, $f \in C(K)$, is non-concave on K if and only if $f(x) \leqq (L_n f)(x)$ for every $x \in K$. Eventually, study the same problem without the second property of the operators.

5880.* *Proposed by Anon, Erewhon-upon-Yarkon*

Let $f(x)$ be a continuous function on $a < x < b$ such that $f'(x)$ exists at each point. Suppose for each x in this interval there exists a $\delta = \delta_x > 0$ such that

$$\frac{f(x+h)-f(x-h)}{2h} = g(x)$$

for all h satisfying $0 < h < \delta$. Prove that $f(x)$ is a quadratic polynomial. (This generalizes a problem in T. M. Flett, *Mathematical Analysis*, where f''' is assumed to exist.)

5881.* *Proposed by D. E. Cooper, Hampton Institute, Virginia*

Let U be a connected open subset of the plane, and let f be a map of U into the plane which is differentiable (in the sense of Walter Rudin, *Principles of Mathematical Analysis*, p. 188). If the Jacobian of f is nonzero at every point of U, must the Jacobian have constant sign?

5882*. *Proposed by E. S. Langford, University of Maine*

Does the set of differentiable functions on the real line have the Riesz Decomposition Property? I.e., if f_1, f_2, and g are positive differentiable functions such that $f_1 + f_2 \geqq g \geqq 0$, can g be written as $g = g_1 + g_2$, where g_1 and g_2 are differentiable functions which satisfy $f_1 \geqq g_1 \geqq 0$ and $f_2 \geqq g_2 \geqq 0$?

5883. *Proposed by Frank Bernhart, Kansas State University*

Given a collection X of subsets of S, no one containing another, let $C(X)$ consist of all minimal subsets of S which intersect every member of X. (1) Show that $C(C(X)) = X$. (2) Characterize collections X such that $C(X) = X$.

SOLUTIONS OF ADVANCED PROBLEMS

Prime Decomposition of $a^b + b^a$

5801 [1971, 549]. *Proposed by Erwin Just, Bronx Community College*

If m and k are arbitrary fixed positive integers and m is odd, prove that (1) there exists a positive integer n such that $m^n + n^m$ contains at least k distinct prime factors, and (2) there exists a positive integer t such that $m^{t+j} + (t+j)^m$ is composite if $j \in \{1, 2, \cdots, k\}$.

Solution by Allen Stenger, Student, Emory University. Designate a set of primes $p_1, p_2, \cdots, p_k$ as follows: First let $p_1 > 2m$. Having selected p_i, let p_{i+1} be of the form

$$sp_1(p_1 - 1)p_2(p_2 - 1) \cdots p_i(p_i - 1) - 1$$

(this is possible by Dirichlet's theorem). If $j < i$, then $p_j < p_i$, so $p_i \nmid (p_j - 1)$. Further

$$p_i - 1 = sp_1(p_1 - 1) \cdots p_{i-1}(p_{i-1} - 1) - 2 \equiv -2 (\operatorname{mod} p_j(p_j - 1)),$$

so $p_j \nmid p_i - 1$ and $(p_i(p_i - 1), p_j(p_j - 1)) = 2$.

(1) Choose n (by the Chinese Remainder Theorem) so that

$$n \equiv -1 (\operatorname{mod} p_i) \text{ for } 1 \leqq i \leqq k \text{ and } n \equiv 0(\operatorname{mod}(p_1 - 1)\cdots(p_k - 1)).$$

Then $m^n + n^m \equiv 1 + (-1)^m \equiv 0 (\operatorname{mod} p_i)$, $1 \leqq i \leqq k$, by Fermat's theorem, since $p_i \nmid m$, and $(p_i - 1) \mid n$, and m is odd. We note that n is even.

(2) Choose (by the Chinese Remainder Theorem) t even, and so large that $m^t + t^m > p_k$, and so that

$$\frac{t}{2} \equiv \frac{n}{2} - i \left(\operatorname{mod} \frac{p_i(p_i - 1)}{2}\right) \qquad 1 \leqq i \leqq k.$$

Then $t + 2i \equiv n \ (\operatorname{mod} p_i(p_i - 1))$, so

$$m^{t+2i} + (t + 2i)^m \equiv m^n + n^m \equiv 0 (\operatorname{mod} p_i), 1 \leqq i \leqq k.$$

Also, $m^{t+2i-1} + (t + 2i - 1)^m$ is even since $t + 2i - 1$ is odd. Hence each of $m^{t+j} + (t + j)^m$, $1 \leqq j \leqq 2k - 1$ is composite, as each is divisible by 2 or by a p_i, and

$$m^{t+j} + (t + j)^m > m^t + t^m > p_k \geqq p_i > 2.$$

Also solved by D. Borwein, Robert Breusch, and the proposer.

Sets with Sequences with Arbitrary Differences

5802 [1971, 678]. *Proposed by J. P. Jones, University of Calgary*

The Cantor set X has the property that for every positive real number, d, X contains points x_0, x_1 such that $d = x_1 - x_0$. More generally, does there exist a set X of measure zero such that for every finite sequence $d_1, d_2, \cdots, d_n$ of positive real numbers, X contains points $x_0, x_1, \cdots, x_n$ such that $d_i = x_i - x_{i-1}$ for $i = 1, 2, \cdots, n$?

Solution by Douglas Lind, Stanford University. The problem is equivalent to finding a set X such that $\bigcap_{j=1}^{n}\{X - (d_1 + \cdots + d_j)\} \neq \varnothing$ for every finite sequence $d_1, \cdots, d_n$ of positive reals. This was settled by R. O. Davies, J. M. Marstrand and S. J. Taylor [Colloq. Math. 7 (1960), 237–243], who gave a simple construction of a closed set X of h-measure 0, where h is a given but arbitrary measure function, such that if $f_1, \cdots, f_n$ are affine maps of the line, then $\bigcap_{j=1}^{n} f_j(X) \neq \varnothing$.

Also solved by D. Borwein & J. M. Borwein, Harold Donnelly, and F. W. Lozier.

Editor's Note. In the paper referred to above, there is a reference to a paper by Erdös and Kakutani, *On a perfect set*, Coll. Math. 4 (1957), p. 195, where the following is proved:

The set $S = \{\sum_{k=2}^{\infty} a_k/k!: 0 \leqq a_k \leqq k - 2, a_k \text{ integral}\}$ has the property that if $x_1 < x_2 < \cdots$

$< x_n$, $x_n - x_{n-1} < \eta_n$ for some $\eta_n > 0$, then there are n elements $y_1, y_2, \cdots, y_n$ of S congruent to $x_1, x_2, \cdots, x_n$; moreover S is perfect and has measure zero.

By suitably placing increasingly magnified copies of S along the real line, we obtain a set satisfying the requirements of the problem.

Rings with Torsion Elements Forming Submodules

5803 [1971, 679]. *Proposed by G. Sabbagh, Yale University*

Let A be a ring with 1 and without zero divisors. It is obvious that, if A is commutative, then the torsion elements of each left A-module E constitute a submodule of E. What other rings have this property?

Note. This problem is the same as 5059 [1963, 1108], as several readers have pointed out. See also 5354 [1967, 96]. Two papers to which we have been referred contain relevant theorems. They are (1) Lawrence Levy, *Torsion-free and divisible modules over a non-integral domain*, Canadian J. Math. 15 (1963), p. 132, ff. See Theorem 1.4, p. 134. (2) Enzo R. Gentile, *Singular submodules and injective hull*, Indagationes Math. 24 (1962), No. 4, p. 430.

Also solved by D. Ž. Djoković & D. J. Fieldhouse, G. J. Janusz, John Kinloch, Israel Kleiner, R. P. Miller, James R. Smith, H. H. Storrer, C. N. Winton, E. T. Wong, and the proposer.

Convex Hull of Unitary Operators on H

5804 [1971, 679]. *Proposed by D. A. Herrero, University of Chicago*

Let $L(H)$ denote the space of all bounded linear operators on the Hilbert space H. Prove that the closed convex hull of the set of all unitary operators is dense in the closed unit hull of $L(H)$.

Solution by E. M. Klein, Northwestern University. Given a contraction A on H (i.e. $\|A\| \leqq 1$) we must approximate A in norm by convex combinations of unitaries. By the solution of problem 107, p. 265 of Halmos, *A Hilbert Space Problem Book* (D. Van Nostrand, 1967) we have $A = \frac{1}{2}(U + V)$, where U and V are maximal partial isometries (a maximal partial isometry is an isometry or a co-isometry). Hence we may assume A is an isometry. By the structure theorem for isometries (Problem 118, *ibid.*) A is the direct sum of a unitary and some copies of the unilateral shift. Therefore we may assume A is the unilateral shift, i.e., A operates on l^2 and $A(x_0, x_1, x_2, \cdots) = (0, x_0, x_1, x_2, \cdots)$.

Let

$$U_n(x_0, x_1, x_2, \cdots) = (\sqrt{1/n}\, x_0, \sqrt{1 - 1/n}\, x_0, x_1, x_2, \cdots)$$

and

$$V_n(x_0, x_1, x_2, \cdots) = (-\sqrt{1/n}\, x_0, \sqrt{1 - 1/n}\, x_0, x_1, x_2, \cdots).$$

U_n and V_n are unitary operators, $\frac{1}{2}(U_n + V_n)(x_0, x_1, x_2, \cdots) = (0, \sqrt{1-1/n}x_0, x_1, x_2, \cdots)$. Hence $\| A - \frac{1}{2}(U_n + V_n) \| = 1 - \sqrt{1 - 1/n}$ which converges to 0. This is the desired result.

Also solved by S. L. Campbell, A. A. Jagers (Netherlands), and the proposer. Seymour Goldberg refers to a paper of Russo and Dye (Duke Math. J. 33 (1966), p. 413 ff.) for a solution to the problem.

On Binary Expansions

5805 [1971, 679]. *Proposed by John Stout, New York, N.Y.*

For any $x \in (0, 1]$, there is exactly one binary expression $x = 0.x_1x_2x_3 \cdots = \sum_{i=1}^{\infty} x_i 2^{-i}$ with an infinite number of ones. Define a function

$$f: (0, 1] \to R^+ \cup \{+\infty\} \text{ by } f(x) = \sum_{i=1}^{\infty} x_i/i.$$

Let C be the inverse image of R^+, and D the inverse image of $\{+\infty\}$.

Are C and D Lebesgue measurable? If so, what are their measures?

Solution by F. V. Meyer, University of Minnesota. Yes; $m(C) = 0$ and $m(D) = 1$. Using the strong law of large numbers, it can be shown (see, e.g., Feller, *Introduction to Probability*, vol. I, p. 195) that almost every number $x \sim \{x_i\}_{i=1}^{\infty}$ is "normal", i.e.,

$$\lim_{n \to \infty} \sum_{i=1}^{n} x_i/n = \tfrac{1}{2}.$$

If x is a fixed normal number, then $(\sum_{i=1}^{n} x_i - n/2) = o(n/2)$; hence, given $\varepsilon > 0$, there is an integer $N > 0$ such that

$$\sum_{i=1}^{n} x_i > (1 - \varepsilon)n/2$$

if $n > N$. Since, for all n

$$\sum_{i=k}^{n} 1/i > (n - k)(1/n),$$

it follows that, if $n > N$ and $n\varepsilon/2$ is an integer, then

$$\sum_{i=n\varepsilon/2}^{n} x_i/i > \sum_{i=(1+2\varepsilon)n/2}^{n} 1/i$$

$$> (1 - 2\varepsilon)(n/2)(1/n) = (1 - 2\varepsilon)/2.$$

Consequently $\sum_{n=1}^{\infty} x_i/i$ diverges.

Also solved by A. N. Al-Hussain, Richard Bagley, D. Borwein & Jon Borwein, R. J. Dickson, Harold Donnelly, John Flaig, G. J. Foschini, Ellen Hertz, A. A. Jagers (Netherlands), H. C. Kranzer,

O. P. Lossers (Netherlands), J. C. Oxtoby, Nicholas Passell, J. H. Roberts, C. C. Rousseau, Tiber Šalát (Czechoslovakia), A. C. Segal, Miha Sharir & Konrad Victor (Israel), Masaaki Shiba (Japan), R. P. Stanley, András Szép (Hungary), and John Mc Cabe & the proposer.

Editorial Note. In her proof, Ellen Hertz cites a result that $\sum_{i=1}^{\infty}(x_i - \frac{1}{2})/i$ converges almost everywhere (Krickeberg, *Probability Theory* p. 109).

Gaussian Integer Power Series

5806 [1971, 679]. *Proposed by Dennis Allen, Jr., Michigan Technological University, Houghton*

Let G denote the ring of Gaussian integers and $G[[x]]$ the ring of formal power series over G. Let $a_1, a_2, \cdots, a_n$ be Gaussian integers, each with positive real part, and let $e_1, \cdots, e_n$ be whole numbers. Suppose

$$\prod_{k=1}^{n}\left[\sum_{j=0}^{\infty}(a_k x)^j\right]^{e_k} = \sum_{j=0}^{\infty} b_j x^j .$$

Does it follow that $b_j \neq 0$ for $0 \leqq j \leqq \min\{e_1, \cdots, e_n\}$?

Solution by P. L. Montgomery, Berkeley, California. The stated result does not follow. Let $n = 2$, $e_1 = e_2 = 4$, $a_1 = 1 + 3i$, $a_2 = 1 - 3i$. Then the coefficient b_2 of x^2 in the expansion of $(1 - a_1x)^{-4}(1 - a_2x)^{-4} = (1 - 2x + 10x^2)^{-4}$ $= 1 - 4(-2x + 10x^2) + 10(-2x + 10x^2)^2 - \cdots$ is 0.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Introduction to Probability Theory and Statistical Inference.* By Harold Larson. Wiley, New York, 1969. xi+388 pp. $10.95. (Telegraphic Review, October 1969.)

Upon casual inspection of this book and its list of contents, my first impression was that this was a routine post-calculus undergraduate text in probability and mathematical statistics. The standard topics of combinatorics, discrete and continu-

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 79 | NUMBER 10

CONTENTS

(Continued on inside cover)

DECEMBER **1972**

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the Monthly. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to Harley Flanders, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

Published by the Association at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

SCHUBERT CALCULUS

S. L. KLEIMAN AND DAN LAKSOV, Massachusetts Institute of Technology

1. Introduction. In 1874, H. Schubert published his celebrated treatise, "*Kalkül der Abzählenden Geometrie*" (Calculus of Enumerative Geometry [**22**]). It dealt with finding the number of points, lines, planes, etc., satisfying certain geometric conditions, an important problem about a hundred years ago. In the book, Schubert drew much from the vast literature on the subject and introduced some far-reaching ideas of his own.

As was often the case in early algebraic geometry, the methods of enumerative geometry were intuitive and rested on a weak foundation. However, the beauty of the subject inspired many mathematicians to develop rigorously the foundational material, such as topological and algebraic intersection theories. This work is of far greater importance than the original enumerative problems.

In a brief article, we can only hope to highlight a rigorous development of the early ideas, but we shall try to illustrate each discussion with an example of lines in 3-space.

Here is a typical enumerative problem: How many lines in 3-space, in general, intersect four given lines? Schubert would specialize the given four lines so as to make the first intersect the second and the third intersect the fourth. In this special case there are obviously two lines intersecting the four: the line joining the two points of intersection and the line of intersection of the two planes—one determined by the first two lines and the other by the second two. Now Schubert's "principle of conservation of number" asserts that there must be two solutions in the general case as well. This principle, which grew out of Poncelet's principle of continuity, is Schubert's most important contribution to the subject.

Our first step will be to make the concept of specializing a line more precise. This we do in section two, where we show more generally that all the d-planes in n-space can in a natural way be made into a manifold. Then we may interpret specialization as moving in a continuous way.

Next, we must analyze the condition that a line L intersect a given line A. This condition means that any two points which determine L and any two points which determine A are dependent and the latter requirement can be conveniently expressed

Steven Kleiman received his Harvard Ph.D. in 1965 under Zariski and Mumford. He was a J.F. Ritt Instructor, then Assistant and Associate Professor at Columbia University before moving to his present Associate Professorship at MIT. He held a NATO Fellowship in 1966–1967 at the IHES-Paris and a Sloan Fellowship in 1968. His main research interest is algebraic geometry, and he is the co-author with Allen Altman of *Introduction to Grothendieck Duality Theory* (Springer Lecture Notes # 146).

Dan Laksov has studied in Norway and France and is presently working on his Doctor's Dissertation under Professor Kleiman at MIT. *Editor.*

in terms of determinants. Section three is devoted to expressing the more general condition that a d-plane in n-space intersect in a prescribed way a given nested sequence (or flag) of linear spaces.

In section four, we interpret and justify the "principle of conservation of number" in the way it was first rigorously done, with the aid of the cohomology theory of manifolds. Then, having defined all our terms, we present the three main theorems of the symbolic formalism, known as Schubert calculus, for solving enumerative problems. We indicate the several different approaches to proving these theorems and give appropriate references in section five.

In section five, we also mention some generalizations, applications, open questions and references pertaining to the material in the other sections. We make no claims of completeness; the choices were made partly out of personal taste. However, we hope that these things will be of interest to some readers and perhaps inspire them to pursue matters further.

2. The Grassmann manifold. The space of n-tuples $(a(1), \cdots, a(n))$ of complex numbers is commonly called **affine** n**-space** and denoted by A^n.

If we try to make sense of the "principle of conservation of number" for configurations in affine space we encounter some difficulties. For example, in section one we found that there are two lines in 3-space which intersect four general lines by specializing the four. However, if we specialize them so that the first intersects the second and the third intersects the fourth but so that the plane of the first two lines is parallel to the plane of the second two, then there will be only one solution. If we specialize the four so that the first intersects the second but the third is parallel to the fourth and the plane of the first two is parallel to the plane of the second two, then there will be no solution. Thus we may obtain 0, 1, or 2 solutions by specializing appropriately. Of course the missing solutions lie "at infinity" and we ought to work in projective space.

A point P of **projective** n**-space** P^n is defined by an $(n+1)$-tuple $(p(0), \cdots, p(n))$ of complex numbers not all zero. The $p(i)$ are called the **coordinates** of P. Another $(n+1)$-tuple $(q(0), \cdots, q(n))$ also defines P if and only if there is a number c satisfying $p(i) = cq(i)$ for $i = 0, \cdots, n$.

Identifying a point $(a(1), \cdots, a(n))$ of A^n with the point $(1, a(1), \cdots, a(n))$ of P^n, we may think of P^r as A^n completed by the points $(0, b(1), \cdots, b(n))$ "at infinity" in P^n. Then, for example, it is not hard to see that two parallel planes, which do not intersect in A^3, will intersect in a line lying "at infinity" in P^3 and that the solutions missing above do lie "at infinity" in this sense.

A **linear space** L in P^n is defined as the set of points $P = (p(0), \cdots, p(n))$ of P^n whose coordinates $p(j)$ satisfy a system of linear equations $\sum_{j=0}^{n} b_{\alpha j} p(j) = 0$ with $\alpha = 1, \cdots, (n-d)$. We say that L is d**-dimensional** if these $(n-d)$ equations are independent, that is if the $(n-d) \times (n+1)$ matrix of coefficients $[b_{\alpha j}]$ has a nonzero $(n-d) \times (n-d)$-minor. By linear algebra, there are then $(d+1)$ points

$P_i = (p_i(0), \cdots, p_i(n))$ in L with $i = 0, \cdots, d$ which span L. Of course, we call L a **line** if $d = 1$, a **plane** if $d = 2$ and a **hyperplane** if $d = (n-1)$. We also call a d-dimensional linear space a **d-plane** for short.

The rest of this section is devoted to representing in a natural way the d-planes in $\boldsymbol{P}^n$ by the points of a certain manifold $G_{d,n}$ lying in a projective space $\boldsymbol{P}^N$ where we put once and for all

$$N = \binom{n+1}{d+1} - 1.$$

For convenience, let us make the following convention. For any $(d+1) \times (n+1)$-matrix $[p_i(j)]$ with $i = 0, \cdots, d$ and $j = 0, \cdots, n$, and any sequence of $(d+1)$ integers $j_0 \cdots j_d$ with $0 \leqq j_\beta \leqq n$, let us denote by $p(j_0 \cdots j_d)$ the determinant of the $(d+1) \times (d+1)$-matrix $[p_i(j_\beta)]$ with $i, \beta = 0, \cdots, d$. Of course, we have the usual formulas:

$$\text{(A)} \qquad \begin{aligned} &p(j_0 \cdots j_d) = 0 \text{ if any two of the } j_\beta \text{ are equal;} \\ &p(j_0 \cdots j_d) = -\,p(j_0 \cdots j_{\beta-1} j_{\beta+1} j_\beta j_{\beta+2} \cdots j_d) \text{ for } \beta = 0, \cdots, d-1. \end{aligned}$$

A function p on the set of all sequences $j_0 \cdots j_d$ with $0 \leqq j_\beta \leqq n$ which satisfies the formulas (A) is called an **alternating function**. It is evident that an alternating function is determined by its values on the subset of sequences $j_0 \cdots j_d$ with $0 \leqq j_0 < \cdots < j_d \leqq n$ and that any function on this subset extends to an alternating function on the whole set. Note that the number of sequences $j_0 \cdots j_d$ with $0 \leqq j_0 < \cdots < j_d \leqq n$ is exactly $(N+1)$.

Fix a d-plane L in $\boldsymbol{P}^n$. Pick $(d+1)$ points $P_i = (p_i(0), \cdots, p_i(n))$ with $i = 0, \cdots, d$ which span L, and form the $(d+1) \times (n+1)$-matrix $[p_i(j)]$. By linear algebra at least one of the $(N+1)$ determinants $p(j_0 \cdots j_d)$ with $0 \leqq j_0 < \cdots < j_d \leqq n$ must be nonzero. So, when ordered lexicographically, these determinants define a point $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $\boldsymbol{P}^N$.

Let $Q_i = (q_i(0), \cdots, q_i(n))$ for $i = 0, \cdots, d$ be another $(d+1)$ points spanning L. Then linear algebra yields a nonsingular $(d+1) \times (d+1)$-matrix C which carries the P_i into the Q_i; in other words, we have $[q_i(j)] = C \cdot [p_i(j)]$ where the dot denotes matrix multiplication. Clearly we then have $q(j_0 \cdots j_d) = \det(C) p(j_0 \cdots j_d)$, where $\det(C)$ denotes the determinant of C. So the points Q_i give rise to the same point of $\boldsymbol{P}^N$ as the points P . Therefore L canonically gives rise to a point of $\boldsymbol{P}^N$. The coordinates $p(j_0 \cdots j_d)$ of this point are called the **Plücker coordinates** of L.

Not every point of $\boldsymbol{P}^N$ arises from some d-plane in $\boldsymbol{P}^n$. In fact, we shall now prove that the Plücker coordinates $p(j_0 \cdots j_d)$ of a d-plane L in $\boldsymbol{P}^n$ satisfy the following quadratic relations:

$$\text{(QR)} \qquad \sum_{\lambda=0}^{d+1} (-1)^\lambda p(j_0 \cdots j_{d-1} k_\lambda) p(k_0 \cdots \check{k}_\lambda \cdots k_{d+1}) = 0,$$

where $j_0 \cdots j_{d-1}$ and $k_0 \cdots k_{d+1}$ are any sequences of integers with $0 \leqq j_\beta, k_\gamma \leqq n$. Here

$\check{k}_\lambda$ means that the integer k_λ has been removed from the sequence and the $p(j_0 \cdots j_d)$ are to be interpreted according to the formulas (A).

Explicitly, we want to establish the relation among determinants,

$$\sum_{\lambda=0}^{d+1} (-1)^\lambda \begin{vmatrix} \vdots & & \vdots & \vdots \\ p_i(j_0) & \cdots & p_i(j_{d-1}) & p_i(k_\lambda) \\ \vdots & & \vdots & \vdots \end{vmatrix} \begin{vmatrix} \cdots & \check{p}_0(k_\lambda) & \cdots \\ & \vdots & \\ \cdots & \check{p}_d(k_\lambda) & \cdots \end{vmatrix} = 0.$$

Expanding the first determinants along their last column, we obtain the relation,

$$\sum_{\lambda=0}^{d+1} (-1)^\lambda \left\{ \sum_{i=0}^{d} (-1)^{d+i} \begin{vmatrix} \vdots & & \vdots \\ \check{p}_i(j_0) & \cdots & \check{p}_i(j_{d-1}) \\ \vdots & & \vdots \end{vmatrix} p_i(k_\lambda) \right\} \begin{vmatrix} \cdots & \check{p}_0(k_\lambda) & \cdots \\ & \vdots & \\ \cdots & \check{p}_d(k_\lambda) & \cdots \end{vmatrix} = 0.$$

Rearranging the terms, we obtain the relation,

$$\sum_{i=0}^{d} (-1)^{d+i} \begin{vmatrix} \vdots & & \vdots \\ \check{p}_i(j_0) & \cdots & \check{p}_i(j_{d-1}) \\ \vdots & & \vdots \end{vmatrix} \left\{ \sum_{\lambda=0}^{d+1} (-1)^\lambda p_i(k_\lambda) \begin{vmatrix} \cdots & \check{p}_0(k_\lambda) & \cdots \\ & \vdots & \\ \cdots & \check{p}_d(k_\lambda) & \cdots \end{vmatrix} \right\} = 0.$$

Now this relation can be obtained by expanding the second determinants in the following relation along the first row:

$$\sum_{i=0}^{d} (-1)^{d+1} \begin{vmatrix} \vdots & & \vdots \\ \check{p}_i(j_0) & \cdots & \check{p}_i(j_{d-1}) \\ \vdots & & \vdots \end{vmatrix} \begin{vmatrix} \cdots & p_i(k_\lambda) & \cdots \\ \cdots & p_0(k_\lambda) & \cdots \\ & \vdots & \\ \cdots & p_d(k_\lambda) & \cdots \end{vmatrix} = 0.$$

However, these second determinants are zero because two rows are equal. Thus the quadratic relations (QR) are satisfied by the Plücker coordinates of a d-plane in $\boldsymbol{P}^n$.

Conversely, any point $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $\boldsymbol{P}^N$ whose coordinates satisfy the quadratic relations (QR) arises from a unique d-plane L in $\boldsymbol{P}^n$. To prove this assertion, we shall simply "solve" the quadratic relations. First, we assume that $p(k_0 \cdots k_d)$ is not zero and show that the $(N+1)$ coordinates $p(j_0 \cdots j_d)$ are already determined by the $[(d+1)(n-d)+1]$ coordinates of the form $p(k_0 \cdots \check{k}_\lambda \cdots k_d j_\alpha)$, that is by the coordinates $p(i_0 \cdots i_d)$ with at most one of $i_0, \cdots, i_d$ not among $k_0, \cdots, k_d$.

Let $j_0 \cdots j_d$ be a sequence of integers of which exactly m are not among the integers $k_0, \cdots, k_d$ and let j_β be one of these m. The quadratic relation (QR) corresponding to the sequences $j_0 \cdots \check{j}_\beta \cdots j_d$ and $k_0 \cdots k_d j_\beta$ obviously yields the equation,

$$p(j_0 \cdots \check{j}_\beta \cdots j_d j_\beta) p(k_0 \cdots k_d) = \sum_{\lambda=0}^{d} (-1)^\lambda p(j_0 \cdots \check{j}_\beta \cdots j_d k_\lambda) p(k_0 \cdots \check{k}_\lambda \cdots k_d j_\beta).$$

Now if k_λ is among $j_0, \cdots, j_d$, then $p(j_0 \cdots \check{j}_\beta \cdots j_d k_\lambda)$ is zero; if k_λ is not among $j_0, \cdots, j_d$ then exactly $(m-1)$ of $j_0, \cdots, \check{j}_\beta, \cdots, j_d, k_\lambda$ are not among $k_0, \cdots, k_d$. Thus if we have $m \geqq 2$, we can express $p(j_0 \cdots j_d)p(k_0 \cdots k_d)$ in terms of the coordinates $p(i_0 \cdots i_d)$ with at most $(m-1)$ of $i_0, \cdots, i_d$ not among $k_0, \cdots, k_d$. Continuing this process of multiplying by $p(k_0 \cdots k_d)$ and of using a quadratic relation, we find we can express $p(j_0 \cdots j_d)p(k_0 \cdots k_d)^{m-1}$ as a polynomial in the coordinates $p(i_0 \cdots i_d)$ with at most one of $i_0, \cdots, i_d$ not among $k_0, \cdots, k_d$. Since we assumed $p(k_0 \cdots k_d) \neq 0$, we have proved our assertion that these $[(d+1)(n-d)+1]$ coordinates determine the others.

Without loss of generality, we assume $p(k_0 \cdots k_d) = 1$. We are going to construct a d-plane L in $\boldsymbol{P}^n$ whose Plücker coordinates are equal to the coordinates $p(j_o \cdots j_d)$ of the given point in $\boldsymbol{P}^N$. For $i = 0, \cdots, d$ and $j = 0, \cdots, n$ put

$$p_i(j) = p(k_0 \cdots k_{i-1} j k_{i+1} \cdots k_d).$$

The vectors $(p_i(0), \cdots, p_i(n))$ for $i = 0, \cdots, d$ are linearly independent because we have $p_i(k_\gamma) = 0$ for $i \neq \gamma$ and $p_i(k_i) = 1$. So, these vectors span a d-plane L in $\boldsymbol{P}^n$. Now the Plücker coordinate $p'(j_0 \cdots j_d)$ of L is defined as the determinant of the matrix $[p_i(j_\beta)]$ with i, $\beta = 0, \cdots, d$. So, if we have $j_\beta = k_\beta$ for $\beta \neq \lambda$, this matrix coincides with the identity matrix outside the λ-th column. Hence we have

$$p'(j_0 \cdots j_d) = p_\lambda(j_\lambda) = p(j_0 \cdots j_d),$$

whenever at most one j_λ of $j_0, \cdots, j_d$ is not among $k_0, \cdots, k_d$. Since we proved above that these coordinates determine the rest, we have $p'(j_0 \cdots j_d) = p(j_0 \cdots j_d)$ for all sequences $j_0 \cdots j_d$. Thus the point $(\cdots, p(j_0 \cdots j_d), \cdots)$ arises from the d-plane L.

Finally, let L' be another d-plane in $\boldsymbol{P}^n$ whose Plücker coordinates define the given point $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $\boldsymbol{P}^N$. Choose $(d+1)$ points $P'_i = (p'_i(0), \cdots, p'_i(n))$ with $i = 0, \cdots, d$ which span L. Then the $(d+1) \times (d+1)$-matrix $[p'_i(k_\gamma)]$ is invertible because its determinant is by hypothesis a nonzero multiple of $p(k_0 \cdots k_d) = 1$. Altering the P'_i by the inverse matrix, we may assume $[p'_i(k_\gamma)]$ is the identity matrix. Then for any sequence $j_0 \cdots j_d$ the determinant $\det[p'_i(j_\beta)]$ is obviously equal to $p(j_0 \cdots j_d)$. Now fix λ and j with $0 \leqq \lambda \leqq d$ and $0 \leqq j \leqq n$, and put $j_\beta = k_\beta$ for $\beta \neq \lambda$ and $j_\lambda = j$. Then $[p'_i(j_\beta)]$ clearly coincides with the identity matrix outside the λ-th column. So we have

$$p'_\lambda(j) = \det[p'_i(j_\beta)] = p(j_0 \cdots j_d) = p_\lambda(j),$$

where the last equation is the definition of $p_\lambda(j)$ made above. Thus we have $P'_\lambda = P_\lambda$ for each λ and so $L' = L$.

We have now reached our goal and proved the following theorem:

THEOREM 1. *There is a natural bijective correspondence between the d-planes in* $\boldsymbol{P}^n$ *and the points of* $\boldsymbol{P}^N$ *with* $N = \binom{n+1}{d+1} - 1$, *whose coordinates satisfy the quadratic relations* (QR).

In the course of the proof, we also established the following result:

PROPOSITION 2. *There is a natural bijective correspondence between the set of points of* $\boldsymbol{P}^N$ *whose coordinates* $p(j_0 \cdots j_d)$ *satisfy the quadratic relations* (QR) *and the requirement* $p(k_0 \cdots k_d) \neq 0$ *and the affine* $(d+1)(n-d)$*-space of* $(d+1)(n+1)$ *matrices* $[p_i(j)]$ *with* $i = 0, \cdots, d$ *and* $j = 0, \cdots, n$ *such that the* $(d+1) \times (d+1)$*-submatrix* $[p_i(k_\gamma)]$ *with* $i, \gamma = 0, \cdots, d$ *is the identity. Moreover, such a matrix* $[p_i(j)]$ *corresponds to the point of* $\boldsymbol{P}^N$ *with coordinates* $p(j_0 \cdots j_d) = \det[p_i(i_\beta)]$ *and such a point* $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $\boldsymbol{P}^N$ *corresponds to the* $(d+1) \times (n+1)$*-matrix with entries*

$$p_i(j) = p(k_0 \cdots k_{i-1} j k_{i+1} \cdots k_d) / p(k_0 \cdots k_d).$$

By virtue of this proposition, the set of points of $\boldsymbol{P}^N$ whose coordinates satisfy the quadratic relations (QR) is covered by $(N+1)$ copies of affine $(d+1)(n-d)$-space, so it is a submanifold of $\boldsymbol{P}^N$ of dimension $(d+1)(n-d)$. It is called the **Grassmann manifold** (of d-planes in n-spaces) and denoted by $G_{d,n}$. In these terms, Theorem (1) says that the d-planes in $\boldsymbol{P}^n$ are represented by the points of the $(d+1)(n-d)$-dimensional Grassmann manifold $G_{d,n}$.

For example, the lines in $\boldsymbol{P}^3$ are represented by the points of the 4-dimensiona Grassmann manifold $G_{1,3}$, which can be described as the points of $\boldsymbol{P}^5$ whose coordinates $p(j_0 j_1)$ satisfy the single quadratic relation,

$$p(01)p(23) - p(02)p(13) + p(03)p(12) = 0.$$

3. Schubert conditions. We are now going to work out a necessary and sufficient determinantal condition for a d-plane in $\boldsymbol{P}^n$ to intersect a given sequence of linear spaces in $\boldsymbol{P}^n$ in a prescribed way.

Let $A_0 \subsetneq A_1 \subsetneq \cdots \subsetneq A_d$ be a strictly increasing sequence (or flag) of $(d+1)$ linear spaces in $\boldsymbol{P}^n$. A d-plane L in $\boldsymbol{P}^n$ is said to satisfy the **Schubert condition** defined by this sequence if $\dim(A_i \cap L) \geqq i$ for all i. The set of all such d-planes L corresponds to a subset of $G_{d,n}$, which is denoted by $\Omega(A_0 \cdots A_d)$.

For example, fix a line A_0 in $\boldsymbol{P}^3$ and take A_1 to be $\boldsymbol{P}^3$ itself. Then the subset $\Omega(A_0 A_1)$ of $G_{1,3}$ represents the set of lines L in $\boldsymbol{P}^3$ satisfying $\dim(L \cap A_0) \geqq 0$ and $\dim(L \cap A_1) \geqq 1$. Since the second condition is automatically satisfied, $\Omega(A_0 A_1)$ represents the set of lines L intersecting A_0.

PROPOSITION 3. *Let* $0 \leqq a_0 < \cdots < a_d \leqq n$ *be a sequence of integers and for* $i = 0, \cdots, d$ *let* A_i *be the* a_i*-dimensional linear space in* $\boldsymbol{P}^n$ *whose points are of the form* $(p(0), \cdots, p(a_i), 0, \cdots, 0)$. *Then* $\Omega(A_0 \cdots A_d)$ *consists exactly of those points* $(\cdots, p(j_0 \cdots j_d), \cdots)$ *in* $G_{d\,n}$ *satisfying* $p(j_0 \cdots j_d) = 0$ *whenever* $j_i > a_i$ *holds for some* i.

Proof. Consider a d-plane L in $\boldsymbol{P}^n$ which satisfies the Schubert condition $\dim(A_i \cap L) \geqq i$ for $i = 0, \cdots, d$. By induction on i, we may clearly pick a point

$P_i = (p_i(0), \cdots, p_i(n))$ in $A_i \cap L$ such that $P_0, \cdots, P_i$ are linearly independent. Then $P_0, \cdots, P_d$ form a basis of L. So, in the construction of section two, L is represented by the point of $G_{d,n}$ with coordinates $p(j_0 \cdots j_d) = \det[p_i(j_\beta)]$. Suppose we have $j_\lambda > a_\lambda$ for a certain λ. Since P_i lies in A_i, we have $p_i(j) = 0$ for $j = (a_i + 1), \cdots, n$, and hence the matrix $[p_i(j_\beta)]$ takes the form,

$$[p_i(j_\beta)] = \begin{array}{cc} & \overbrace{\hphantom{0000}}^{(d-\lambda+1)} \\ \left[\begin{array}{c} * \\ * \end{array}\right. & \left.\begin{array}{c} \boxed{\;0\;} \\ * \end{array}\right]\begin{array}{l} \}\,\lambda \\ \end{array} \end{array}$$

It is now easy to see that $p(j_0 \cdots j_d) = 0$ either by (Laplace) expansion of the determinant along the last $(d - \lambda + 1)$ columns or by induction on $(d - \lambda + 1)$, the cases $(d - \lambda + 1) = 1$ and $(d - \lambda + 1) = 2$ being clear.

Conversely consider a point $(\cdots, p(j_0 \cdots j_d), \cdots)$ on $G_{d,n}$ satisfying $p(j_0 \cdots j_d) = 0$ whenever $j_i > a_i$ holds for some i. Choose a nonzero coordinate $p(k_0 \cdots k_d)$ which maximizes the sum $\Sigma_{\gamma=0}^{d} k_\gamma$. Replacing each $p(j_0 \cdots j_d)$ by $p(j_0 \cdots j_d)/p(k_0 \cdots k_d)$, we may assume $p(k_0 \cdots k_d) = 1$. Now, in section two, we saw that the point $(\cdots, p(j_0 \cdots j_d), \cdots)$ represents the d-plane L spanned by the points $P_i = (p_i(0), \cdots, p_i(n))$ with $p_i(j) = p(\cdots k_{i-1} j k_{i+1} \cdots)$ for $j = 0, \cdots, n$ and for $i = 0, \cdots, d$.

Fix $j > a_i$, we shall show that $p_i(j)$ is zero. Since $p(k_0 \cdots k_d)$ is not zero, we have $k_i \leqq a_i$ and so $k_i < j$. Consequently, the sum $\Sigma_{\gamma=0}^{d} k_\gamma$ is strictly less than the sum $(j + \Sigma_{\gamma \neq i} k_\gamma)$. Hence $p_i(j) = p(\cdots k_{i-1} j k_{i+1} \cdots)$ is zero by the maximality of $\Sigma_{\gamma=0}^{d} k_\gamma$.

Therefore P_i lies in A_i. Hence the $(i + 1)$ linearly independent points $P_0, \cdots, P_i$ lie in $(A_i \cap L)$. So L satisfies the Schubert condition $\dim(A_i \cap L) \geqq i$ for $i = 0, \cdots, d$. Thus $(\cdots, p(j_0 \cdots j_d), \cdots)$ lies in $\Omega(A_0 \cdots A_d)$.

PROPOSITION 4. *Let $A_0 \subsetneqq \cdots \subsetneqq A_d$ and $B_0 \subsetneqq \cdots \subsetneqq B_d$ be two strictly increasing sequences of linear spaces in $\boldsymbol{P}^n$ and assume $\dim(A_i) = \dim(B_i)$ for $i = 0, \cdots, d$. Then there is an invertible linear transformation of $\boldsymbol{P}^N$ into itself which carries $G_{d,n}$ into itself and $\Omega(B_0 \cdots B_d)$ into $\Omega(A_0 \cdots A_d)$.*

Proof. Since we have $\dim(A_i) = \dim(B_i)$ for each i, there obviously is an invertible $(n + 1) \times (n + 1)$-matrix $[a_{ij}]$ such that the linear transformation T of $\boldsymbol{P}^n$ into itself defined by the formula

$$T(p(0), \cdots, p(n)) = \left(\sum_{i=0}^{n} p(i) a_{i0}, \cdots, \sum_{i=0}^{n} p(i) a_{in} \right)$$

carries B_i onto A_i for each i. Clearly, T carries a d-plane L in $\boldsymbol{P}^n$ into another one $T(L)$, and if L satisfies the Schubert condition $\dim(B_i \cap L) \geqq i$ for all i, then $T(L)$

satisfies the Schubert condition $\dim(A_i \cap T(L)) \geqq i$ for all i because we have $T(B_i) = A_i$.

Choose $(d+1)$ points $P_i = (p_i(0), \cdots, p_i(n))$ with $i = 0, \cdots, d$ which span L. Then the $(d+1)$ points $T(P_i)$ span $T(L)$. Now, $T(P_i)$ is of the form $(q_i(0), \cdots, q_i(n))$ with

$$q_i(j) = \sum_{\alpha=0}^{n} p(\alpha) a_{\alpha j} \text{ for } j = 0, \cdots, n,$$

and a straightforward computation shows that the Plücker coordinates $q(j_0 \cdots j_d) = \det[q_i(j_\beta)]$ of $T(L)$ are certain fixed linear combinations of the Plücker coordinates $p(j_0 \cdots j_d) = \det[p_i(j_\beta)]$ of L.

In other words, there is a linear transformation $\Lambda[a_{ij}]$ of $\boldsymbol{P}^N$ into itself which carries $G_{d,n}$ into itself and $\Omega(B_0 \cdots B_d)$ into $\Omega(A_0 \cdots A_d)$. Since $[a_{ij}]$ is nonsingular, it is evident that $\Lambda[a_{ij}]$ is invertible and $\Lambda([a_{ij}]^{-1})$ is its inverse.

COROLLARY 5. *Let $B_0 \subsetneqq \cdots \subsetneqq B_d$ be a strictly increasing sequence of linear spaces in $\boldsymbol{P}^n$. Then $\Omega(B_0 \cdots B_d)$ consists of those points in $G_{d,n}$ whose coordinates $q(j_0 \cdots j_d)$ satisfy certain linear equations; in other words, $\Omega(B_0 \cdots B_d)$ is the intersection of $G_{d,n}$ and a certain linear space in $\boldsymbol{P}^N$. Moreover, the linear space is a hyperplane if and only if we have $\dim(B_0) = (n-d-1)$ and $\dim(B_i) = (n-d+i)$ for $i = 1, \cdots, d$.*

Proof. For $i = 0, \cdots, d$ put $a_i = \dim(B_i)$ and let A_i be the a_i-dimensional linear space in $\boldsymbol{P}^n$ whose points are of the form $(p(0), \cdots, p(a_i), 0, \cdots, 0)$. By Proposition 4, there is a linear transformation S of $\boldsymbol{P}^N$ into itself such that a point P of $G_{d,n}$ lies in $\Omega(B_0 \cdots B_d)$ if and only if $S(P)$ lies in $\Omega(A_0 \cdots A_d)$. By virtue of Proposition 3, $S(P)$ lies in $\Omega(A_0 \cdots A_d)$ if and only if each of its coordinates $q(j_0 \cdots j_d)$ is zero whenever $j_i > a_i$ holds for some i. Since each $q(j_0 \cdots j_d)$ is a certain linear combination of the coordinates $p(j_0 \cdots j_d)$ of P, we conclude that P lies in $\Omega(B_0 \cdots B_d)$ if and only if the $p(j_0 \cdots j_d)$ satisfy certain linear equations. Moreover, the number of linearly independent equations is obviously the number of sequences $j_0 \cdots j_d$, such that $j_i > a_i$ holds for some i, and it is evident that there is only one such sequence if and only if we have $a_0 = (n-d-1)$ and $a_i = (n-d+i)$ for $i = 1, \cdots, d$. Thus, the Corollary is proved.

We are now in a good position to determine the number of lines L in $\boldsymbol{P}^3$ which (simultaneously) intersect four given lines L_1, L_2, L_3, L_4. In section two, we saw that the lines L are represented by the set $G_{1,3}$ of points $(p(01), p(02), p(03), p(12), p(13), p(23))$ of $\boldsymbol{P}^5$ which satisfy the single quadratic relation

$$p(01)p(23) - p(02)p(13) + p(03)p(12) = 0.$$

At the beginning of this section, we noted that the lines L intersecting a given line A are represented by the points of the subset $\Omega(A\boldsymbol{P}^3)$ of G_{13}; hence, the lines L intersecting the four given lines L_1, L_2, L_3, L_4 are represented by the points of the intersection

$$Q = \bigcap_{i=1}^{4} \Omega(L_i\boldsymbol{P}^3).$$

Now, by Corollary 5, for each i we have $\Omega(L_i\boldsymbol{P}^3) = G_{1,3} \cap H_i$ for a suitable hyperplane H_i of $\boldsymbol{P}^5$. Put $M = \bigcap_{i=1}^{4} H_i$; then we have $Q = G_{1,3} \cap M$. If the H_i are linearly independent, then M is a line. Then, by using the quadratic relation defining $G_{1,3}$ to express Q as the zeros of a certain quadratic polynomial in a parameter of M, it is easy to see that Q consists of two points, which may coincide. (They coincide exactly when M is tangent to $G_{1,3}$.) If the H_i are linearly dependent, then M is a linear space of dimension two or more and it is easy to see that Q must be infinite. Thus, the number of lines L which intersect L_1, L_2, L_3, and L_4 is either infinity or two or one (counted twice).

It is not hard to choose the lines L_1, L_2, L_3. L_4 in such a way that Q consists of only one point. Consequently, the "principle of conservation of number" will not be valid unless multiplicities are taken into account. For example, take L_1, L_2, and L_3 to be three skew lines. Fix a point P_1 on L_1. Let π_2 be the plane of P_1 and L_2 and let π_3 be the plane of P_1 and L_3. Since L_2 and L_3 do not intersect, the planes π_2 and π_3 are distinct. Take L_4 to be the line of intersection of these two planes. Then L_4 passes through P_1 and it intersects L_2 in a point P_2 and L_3 in a point P_3. The points P_1, P_2, and P_3 are distinct because the lines L_1, L_2 and L_3 are skew, so any two of the points determine L_4. Now let L be any line intersecting L_1, L_2, L_3 and L_4. If L passes through P_2 and P_3, then L coincides with L_4 because P_2 and P_3 determine L_4. Suppose L does not pass through P_2. Since L intersects L_2 and L_4, it must then lie in the plane of L_2 and L_4, which is π_2. So L passes through the point of intersection of π_2 and L_1, which is P_1. Similarly L must also pass through P_3. Then L coincides with L_4 because P_1 and P_3 determine L_4. Thus L_4 is the only line intersecting L_1, L_2, L_3, and L_4.

In the above example we saw that for any three skew lines L_1, L_2, L_3 in $\boldsymbol{P}^3$ there is a unique line which passes through a given point P_1 of L_1 and intersects L_2 and L_3. Hence, if we had chosen L_4 to be L_1 itself, then there would be an infinite number of lines intersecting L_1, L_2, L_3, and L_4, one for each point of L_1. Of course, the number of lines intersecting four given lines is also infinite if the four all pass through the same point or if they all lie in the same plane.

Since an infinite number of solutions do appear in some special cases of an enumerative problem, the "principle of conservation of number" must be stated in the following way: If the number of solutions is finite in a given special case, then the number of solutions is the same in the general case as well, multiplicities, of course, being taken into account. In some problems, as in determining the lines in 3-space which intersect three given lines, the number of solutions is infinite. In these problems, the "principle of conservation of number" does not strictly apply. However, as Schubert himself realized, something is conserved under specialization. In the next section, we shall see that what is conserved is a cohomology class.

4. The Schubert calculus. In this section we explain the symbolic formalism, known as Schubert calculus, for solving enumerative problems. The foundational material here is far deeper than before and the main proofs are far more difficult, so we shall not go into them. However, we shall indicate the various ways to approach them and give references in the next section.

We shall base our development upon algebraic topology. In section two, we saw that $G_{d,n}$ is a complex manifold of dimension $(d+1)(n-d)$. From algebraic topology, we know that the cohomology group with the integers as coefficients $H^i(G_{d,n};\boldsymbol{Z})$ is zero when i is not in the interval $[0,2(d+1)(n-d)]$ and that the direct sum

$$H^*(G_{d,n};\boldsymbol{Z}) = \bigoplus_i H^i(G_{d,n};\boldsymbol{Z})$$

is a graded ring under cup-product. Moreover, $G_{d,n}$ is oriented, so there is a natural isomorphism of the $2(d+1)(n-d)$-th cohomology group with $\boldsymbol{Z}$; the image in $\boldsymbol{Z}$ of an element u is called the **degree** of u and denoted by $\deg(u)$.

A harder result is that we can assign a natural cohomology class (that is, an element of $H^*(G_{d,n};\boldsymbol{Z})$) to each subset of $G_{d,n}$ defined by a system of polynomial equations. Such a subset is called a **subvariety** of $G_{d,n}$. If two subvarieties are members of the same continuous system of subvarieties, then both are assigned the same cohomology class. (Intuitively, the two are homotopic.)

The subsets $\Omega(A_0\cdots A_d)$ are subvarieties of $G_{d,n}$ by Corollary 5; they are called **Schubert varieties** and their cohomology classes are called **Schubert cycles**. We are now going to prove that the cohomology class of $\Omega(A_0\cdots A_d)$ depends only on the integers $a_i=\dim(A_i)$ for $i=0,\cdots,d$. Indeed, consider the continuous system of subvarieties $(\Lambda M)\Omega(A_0\cdots A_d)$ parametrized by the nonsingular $(n+1)\times(n+1)$-matrices M, where ΛM denotes the linear transformation of $\boldsymbol{P}^N$ into itself induced by the matrix M, (see the proof of Proposition 4). This system clearly includes $\Omega(A_0\cdots A_d)$ and by Proposition 4 it includes every subvariety $\Omega(B_0\cdots B_d)$ with $\dim(B_i)=a_i$ for $i=0,\cdots,d$. Since all the subvarieties in a continuous system are assigned the same cohomology class, the cohomology class of $\Omega(A_0\cdots A_d)$ depends only on the a_i. We are now justified in denoting this Schubert cycle by $\Omega(a_0\cdots a_d)$.

Perhaps the most important result in the theory of cohomology classes is this: When several subvarieties intersect properly in a finite set of points, then the number of points, counted with multiplicity, is equal to the degree of the product of the corresponding cohomology classes. Roughly put, the theorem holds because passing to cohomology classes turns intersection into cup-product. For example, suppose each subvariety represents the d-planes in $\boldsymbol{P}^n$ which satisfy certain geometric conditions. Then the number of d-planes which simultaneously satisfy all the conditions, multiplicities being taken into account, can be determined by formally computing with the corresponding cohomology classes. Since the cohomology classes all remain the same when the subvarieties vary in a continuous system, this number will remain

constant when the geometric conditions are varied (or specialized) in a continuous way. This conclusion is an interpretation of Schubert's "principle of conservation of number."

We now state the first main theorem of Schubert calculus. It asserts that the Schubert cycles completely determine the cohomology of $G_{d,n}$.

THEOREM (The basis theorem). *Considered additively $H^*(G_{d,n}; \mathbf{Z})$ is a free abelian group and the Schubert cycles $\Omega(a_0 \cdots a_d)$ form a basis.*

By construction, the cohomology class of a subvariety X of $G_{d\,n}$ lies in $H^{2p}(G_{d,n}; \mathbf{Z})$ when X is irreducible of dimension $[(d+1)(n-d)-p]$. Irreducibility means that X is not the union of two smaller subvarieties in a nontrivial way. The dimension of X is then r if an open subset of X is canonically a manifold of dimension r.

We now prove that $\Omega(A_0 \cdots A_d)$ is irreducible of dimension $\Sigma_{i=0}^{d}(a_i - i)$ with $a_i = \dim(A_i)$. First, suppose A_i consists of the points $(p(0), \cdots, p(n))$ with $p(j) = 0$ when $j > a_i$ and consider the space S of all $(d+1) \times (n+1)$-matrices $[p_i(j)]$ with $p_i(j) = 0$ when $j > a_i$ for $i = 0, \cdots, d$. Let S_0 be the open subset of S of matrices whose maximal minors $p(j_0 \cdots j_d) = \det[p_\alpha(j_\beta)]$ are not all zero. In the course of proving Proposition 3 we saw that sending a matrix $[p_i(j)]$ to the point $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $\mathbf{P}^N$ defines a map π of S_0 onto $\Omega(A_0 \cdots A_d)$. Since S is an affine space, it follows by an elementary argument that S_0 is irreducible and consequently that $\Omega(A_0 \cdots A_d)$ is irreducible. Now, let S_1 be the subset of S of matrices $[p_i(j)]$ whose submatrix $[p_i(a_j)]$ is the $(d+1) \times (d+1)$ identity. Then S_1 lies in S_0 and as we saw when proving Proposition 3, $\pi(S_1)$ is the open subset of $\Omega(A_0 \cdots A_d)$ of points $(\cdots, p(j_0 \cdots j_d), \cdots)$ with $p(a_0 \cdots a_d) \neq 0$. However, Proposition 2 implies that π induces an analytic isomorphism of S_1 with $\pi(S_1)$. Since S_1 is obviously an affine space of dimension $\Sigma_{i=0}^{d}(a_i - i)$, the dimension of $\Omega(A_0 \cdots A_d)$ is therefore this number.

We may now rephrase the basis theorem in the following way:

THEOREM (The basis theorem). *Each even dimensional integral cohomology group $H^{2p}(G_{d,n}; \mathbf{Z})$ is a free abelian group and the Schubert cycles $\Omega(a_0 \cdots a_d)$ with $[(d+1)(n-d) - \Sigma_{i=0}^{d}(a_i - i)] = p$ form a basis. Each odd dimensional group is zero.*

For example, consider the Grassmann manifold $G_{0,n}$ of points in $\mathbf{P}^n$. The Plücker coordinates of a point are obviously its ordinary coordinates; hence, we have $G_{0,d} = \mathbf{P}^n$ and $\Omega(A_0) = A_0$. Now, the basis theorem says that $H^{2p}(\mathbf{P}^n; \mathbf{Z})$ for $0 \leqq p \leqq n$ is a free cyclic group generated by the class $\Omega(n-p)$ of an $(n-p)$-dimensional linear space. The other groups are zero.

For a second example, consider the Grassmann manifold $G_{1,3}$ of lines in $\mathbf{P}^3$. Here, the basis theorem says that there are exactly five nonzero cohomology groups:

the middle one $H^4(G_{1,3}; \mathbf{Z})$ is free abelian on two generators $\Omega(0.3)$ and $\Omega(1.2)$ and the others $H^{2p}(G_{1,3}; \mathbf{Z})$ for $p = 0,1,3,4$ are free cyclic on generators respectively $\Omega(2.3)$, $\Omega(1.3)$, $\Omega(0.2)$, $\Omega(0.1)$. Moreover, it is evident that $\Omega(0.1)$ is the class of a point, that $\Omega(2.3)$ is the class of $G_{1,3}$ and that in view of Corollary 5, $\Omega(1.3)$ is the class of a hyperplane section.

The following proposition complements the basis theorem with some very useful information.

PROPOSITION. *The basis* $\{\cdots,\Omega(a_0 \cdots a_d),\cdots\}$ *of the group* $H^{2p}(G_{d,n}; \mathbf{Z})$ *and the basis* $\{\cdots,\Omega(n-a_d,\cdots,n-a_0),\cdots\}$ *of the group* $H^{2[(d+1)(n-d)-p]}(G_{d,n}; \mathbf{Z})$ *are dual under the pairing* $v,w \mapsto \deg(v \cdot w)$ *of Poincaré duality.*

In other words, the proposition says that an arbitrary element v of $H^{2p}(G_{d,n}; \mathbf{Z})$ can be written uniquely in the form

$$v = \sum \delta(n-a_d,\cdots,n-a_0)\Omega(a_0 \cdots a_d),$$

where the integers $\delta(n-a_d,\cdots,n-a_0)$ can be found by using the formula

$$\delta(n-a_d,\cdots,n-a_0) = \deg(v \cdot \Omega(n-a_d,\cdots,n-a_0)).$$

In particular, if v is the cohomology class of an irreducible subvariety X of $G_{d,n}$, then each integer $\delta(n-a_d,\cdots,n-a_0)$ is nonnegative because it is the number of points with multiplicity in the intersection of X and $\Omega(B_0 \cdots B_d)$ for suitably chosen linear spaces B_i. Schubert called these integers the degrees (*Gradzahlen*) of X.

Let Y be an irreducible subvariety of $G_{d,n}$ of dimension p and let the integers $\varepsilon(a_0 \cdots a_d)$ be its degrees. If the intersection $X \cap Y$ is a finite set of points, then the number $i(X \cap Y)$ of points counted with multiplicity is, as we know, the degree of the product of

$$\sum\delta(n-a_d,\cdots,n-a_0)\Omega(a_0 \cdots a_d) \text{ and } \sum\varepsilon(a_0 \cdots a_d)\Omega(n-a_d,\cdots,n-a_0).$$

Therefore, by the proposition we have

$$i(X \cap Y) = \sum\delta(n-a_d,\cdots,n-a_0)\varepsilon(a_0 \cdots a_d).$$

This formula constitutes a generalization of Bézout's theorem. Bézout's theorem deals with the case $G_{0,n} = \mathbf{P}^n$. We saw above that the cohomology class v of an $(n-p)$-dimensional irreducible subvariety X of $\mathbf{P}^n$ is of the form $v = \delta(p)\Omega(n-p)$ and by the proposition $\delta(p)$ is the number of points with multiplicity in the intersection of X and a suitably chosen p-dimensional linear space. Thus $\delta(p)$ is the degree of X in the usual sense. Let Y be a p-dimensional irreducible subvariety of $\mathbf{P}^n$ and let $\varepsilon(n-p)$ be its degree. Suppose X and Y intersect in a finite set of points. Then the formula above becomes $i(X \cap Y) = \delta(p)\varepsilon(n-p)$; in other words, the number of points counted with multiplicity in $X \cap Y$ is the product of the degree of X and the degree of Y. This result is known as Bézout's theorem.

The basis theorem implies that the product of any two Schubert cycles can be uniquely expressed as a linear combination of other Schubert cycles with integers as coefficients. The second and third main theorems allow us to compute such expressions explicitly. The second expresses an arbitrary Schubert cycle as a determinant in the following $(n-d+1)$ special Schubert cycles:

$$\sigma(h) = \Omega(h, n-d+1, \cdots, n) \text{ for } h = 0, \cdots, (n-d).$$

THEOREM (The determinantal formula). *For all sequences of integers* $0 \leqq a_0 < \cdots < a_d \leqq n$ *the following formula holds in the cohomology ring* $H^*(G_{d,n}; \mathbf{Z})$:

$$\Omega(a_0 \cdots a_d) = \begin{vmatrix} \sigma(a_0) & \cdots & \sigma(a_0 - d) \\ \vdots & & \vdots \\ \sigma(a_d) & \cdots & \sigma(a_d - d) \end{vmatrix}$$

where we agree to put $\sigma(h) = 0$ *for* $h \notin [0, (n-d)]$.

This theorem, together with the basis theorem, implies that the special Schubert cycles generate the cohomology ring as a $\mathbf{Z}$-algebra. Moreover, it reduces the problem of determining the product of two arbitrary Schubert cycles to the case where one (or for that matter, each) is a special Schubert cycle. This case is handled by the third main theorem, which follows.

THEOREM (Pieri's formula). *For all sequences of integers* $0 \leqq a_0 < \cdots < a_d \leqq n$ *and for* $h = 0, \cdots, (n-d)$, *the following formula holds in the cohomology ring* $H^*(G_{d,n}; \mathbf{Z})$:

$$\Omega(a_0 \cdots a_d) \cdot \sigma(h) = \Sigma\Omega(b_0 \cdots b_d),$$

where the sum ranges over all sequences of integers $b_0 < \cdots < b_d$ *satisfying* $0 \leqq b_0 \leqq a_0 < b_1 \leqq a_1 < \cdots < b_d \leqq a_d$ *and* $\Sigma_{i=0}^{d} b_i = \Sigma_{i=0}^{d} a_i - (n-d-h)$.

Let us use these results to determine the number of lines L in $\mathbf{P}^3$ which (simultaneously) intersect four given lines L_1, L_2, L_3, L_4. In section three, we saw that such lines L are represented by the points of the intersection

$$Q = \bigcap_{i=1}^{4} \Omega(L_i, \mathbf{P}^3).$$

So, we want to compute the degree of $\Omega(1.3)^4$. By definition we have $\Omega(1.3) = \sigma(1)$ and Pieri's formula gives $\Omega(1.3) \cdot \sigma(1) = \Sigma\Omega(b_0 \cdot b_1)$ with $0 \leqq b_0 \leqq 1 < b_1 \leqq 3$ and $b_0 + b_1 = 3$. Hence we obtain $\Omega(1.3)^2 = \Omega(0.3) + \Omega(1.2)$. Now, the proposition yields $\Omega(0.3)^2 = 0$, $\Omega(1.2)^2 = 0$ and $\deg(\Omega(0.3) \cdot \Omega(1.2)) = 1$. Hence we find $\deg(\Omega(1.3)^4) = 2$. Alternately, a second application of Pieri's formula yields $\Omega(1.3)^3 = 2\Omega(0.2)$ and a third yields $\Omega(1.3)^4 = 2\Omega(0.1)$. Since $\Omega(0.1)$ is the class of a single point, its degree

is one. Thus, we again find $\deg(\Omega(1.3)^4) = 2$. Therefore, if Q is a finite set of points, then the number of points with multiplicity in Q is two. Thus the number of lines is either infinity or two or one (counted twice).

In the preceding example we obtained the formula $\Omega(1.3)^3 = 2\Omega(0.2)$. Since the various subvarieties in a continuous system are all assigned the same cohomology class, this formula suggests that the set of lines which simultaneously intersect three skew lines can be continuously deformed into the union of two sets of lines which lie in a plane and pass through a fixed point. In fact, we shall now see that this is the case.

Specialize the three lines L_1, L_2, L_3 so that L_1 and L_2 intersect in a point P and so that L_3 intersects the plane F of L_1 and L_2 in a point Q not equal to P. Then a line intersecting L_1 and L_2 must either lie in F or pass through P, and conversely a line lying in F or passing through P intersects L_1 and L_2. So a line intersecting L_1, L_2 and L_3 must either lie in F and pass through Q or pass through P and lie in the plane F' of P and L_3, and conversely a line lying in F and passing through Q or lying in F' and passing through P intersects L_1, L_2 and L_3. In other words, we have

$$\bigcap_{i=1}^{3} \Omega(L_i \cdot \boldsymbol{P}^3) = \Omega(Q \cdot F) + \Omega(P \cdot F').$$

When the subvarieties of $G_{d,n}$ defined by more general geometric conditions are considered, the power of the calculus becomes staggering. Schubert's book contains many examples and we now give two.

Let us compute the number of lines L in $\boldsymbol{P}^3$ which simultaneously intersect four given curves C_1, C_2, C_3, C_4. Let $c_i \in H^4(\boldsymbol{P}^3; \boldsymbol{Z})$ be the cohomology class of C_i and ℓ the class of a line. We have $c_i = \delta_i \ell$, where δ_i is the degree of C_i, (see the discussion of Bezout's theorem after the proposition). So it is not surprising (and is justified below) that the lines L which intersect a given C_i are represented by the points of a subvariety X_i of $G_{1,3}$ and that the cohomology class x_i of X_i is of the form $x_i = \delta_i \Omega(1.3)$. Hence we have

$$x_1 x_2 x_3 x_4 = 2\delta_1 \delta_2 \delta_3 \delta_4 \Omega(0.1)$$

in view of the computations in the example above. So when the number of lines intersecting C_1, C_2, C_3, C_4 is finite and multiple solutions are taken into account, the number of lines is $2\delta_1\delta_2\delta_3\delta_4$. This result is indicated geometrically by specializing each C_i so that it becomes a union of δ_i lines, then the number of lines (simultaneously) intersecting C_1, C_2, C_3, C_4 is obviously $\delta_1\delta_2\delta_3\delta_4$ times the number of lines intersecting four lines and the latter number, we know, is 2.

To analyze each X_i rigorously, we need to consider the subset Z of the product $\boldsymbol{P}^3 \times G_{1,3}$ consisting of the pairs (P, Q) such that the point P of $\boldsymbol{P}^3$ lies on the line represented by Q. With a certain amount of elementary computations like those in sections one and two, one can show that Z is a complex manifold of dimension 5 which can be described by a system of (bihomogenous quadratic) polynomial

equations. Let $p: \boldsymbol{P}^3 \times G_{1,3} \to \boldsymbol{P}^3$ and $q: \boldsymbol{P}^3 \times G_{1,3} \to G_{1,3}$ be the projections. Then we clearly have $X_i = q(Z \cap p^{-1}C_i)$ set-theoretically and it is easy to show that $x_i = q_*(z \cdot p^*c_i)$, where z is the cohomology class of Z, where p^* is the natural operation on cohomology induced by p and where q_* is the Poincaré dual of q^*. Similarly we have $\Omega(1.3) = q_*(z \cdot p^*\ell)$. Consequently the relation $c_i = \delta_i \ell$ implies the relation $x_i = \delta_i q_*(z \cdot p^*\ell) = \delta_i\Omega(1.3)$ as asserted.

Finally, we sketch a proof that two quadrics in $\boldsymbol{P}^4$ have, in general, sixteen lines in common. A quadric Q in $\boldsymbol{P}^n$ is defined as the set of zeros of a single homogeneous polynomial F of degree two and the $m = \binom{n+2}{2}$ coefficients of F may be used to represent Q by a point q of $\boldsymbol{P}^{m-1}$. First, we observe that the lines L in $\boldsymbol{P}^4$ which lie on a general quadric are represented by the points ℓ of a 3-dimensional irreducible subvariety of $G_{1,4}$. Indeed, let W be the subset of $\boldsymbol{P}^{14} \times G_{1,4}$ consisting of the pairs (q,ℓ) where q represents a quadric Q in $\boldsymbol{P}^4$ and ℓ represents a line L lying in Q. Let $p: W \to \boldsymbol{P}^{14}$ and $r: W \to G_{1,4}$ be the projections. A fiber of r represents the quadrics Q which contain a given line L. Let F_1, F_2, F_3 be independent homogeneous linear equations defining L. Then the polynomial F defining Q is obviously of the form

$$F = G_1F_1 + G_2F_2 + G_3F_3,$$

where G_i is a suitable homogeneous linear equation. Hence all such polynomials F form a vector space of dimension $(5 + 4 + 3) = 12$, so the fiber of r is $\boldsymbol{P}^{11}$. Therefore W is an irreducible subvariety of dimension $[11 + \dim(G_{1,4})] = 17$. A general fiber of p, which represents the lines lying on a general quadric Q, is therefore irreducible of dimension $(17 - 14) = 3$.

Let Q be a general quadric in $\boldsymbol{P}^4$, let X be the 3-dimensional irreducible subvariety of $G_{1,4}$ representing the lines lying in Q, and let x be the cohomology class of X. By the basis theorem, we have $x = \lambda\Omega(0.4) + \mu\Omega(1.3)$ and by the proposition, we have $\lambda = \deg(x \cdot \Omega(0.4))$ and $\mu = \deg(x \cdot \Omega(1.3))$. Now, no line lying in Q can pass through a point P of $\boldsymbol{P}^4$ not in Q. Hence $X \cap \Omega(P, \boldsymbol{P}^4)$ is empty, so we have $\lambda = 0$. On the other hand, a general 3-dimensional linear space A_1 intersects Q in a quadric Q_1 in this copy of $\boldsymbol{P}^3$ and exactly four lines lying in Q_1 meet a general line A_0 lying in A_1 because A_0 intersects Q_1 in two distinct points and therefore meets a line of each ruling at each point. Hence $X \cap \Omega(A_0 \cdot A_1)$ consists of four points, so we have $\mu = 4$. Let Q' be another general quadric in $\boldsymbol{P}^4$. Then the number of lines common to Q and Q', multiplicities being taken into account, is therefore equal to $\deg(x^2) = 4^2\deg(\Omega(1.3)^2) = 16$.

5. Some comments, references and open questions. — Nearly everything discussed so far remains valid in characteristic p. The cohomology theory used in section four has been completely algebrized, and the material of sections two and three generalizes virtually without change over any ground field. In what follows, we shall work over an arbitrary ground field k and discuss restrictions on k as needed.

The work of Hodge and Pedoe **[8]** is by far the most complete reference. Their

treatment is purely algebraic and largely independent of the blanket assumption of characteristic zero.

Concerning section two.—Proceeding as in the first part of the proof of Theorem 1, however, using (Laplace) expansion of the determinants along several columns, one proves that the Plücker coordinates of a d-plane in $\boldsymbol{P}^n$ satisfy more quadratic relations, namely

$$\sum_{\sigma} \operatorname{sgn}(\sigma)\, p(i_0 \cdots i_{\lambda-1}\sigma i_\lambda \cdots \sigma i_d)p(\sigma j_0 \cdots \sigma j_\lambda j_{\lambda+1} \cdots j_d) = 0,$$

where the sum ranges over all permutations σ of $(i_\lambda \cdots i_d j_0 \cdots j_\lambda)$ such that $\sigma i_\lambda < \cdots < \sigma i_d$ and $\sigma j_0 < \cdots < \sigma j_\lambda$. The quadratic relations (QR) occur when we take $\lambda = d$.

For each sequence of integers $i_0 \cdots i_d$ satisfying $0 \leqq i_0 < i_1 < \cdots < i_d \leqq n$ take an indeterminate $X(i_0 \cdots i_d)$ and then, by using the formulas (A), define $X(i_0 \cdots i_d)$ for any sequence of integers $i_0 \cdots i_d$ satisfying $0 \leqq i_j \leqq n$ for $j = 0, \cdots, d$. In these terms, we can now say that $G_{d,n}$ is contained in the set of zeros of all the homogeneous quadratic polynomials of the form

$$\text{(QP)} \qquad \sum_{\sigma} \operatorname{sgn}(\sigma) X(i_0 \cdots i_{\lambda-1}\sigma i_\lambda \cdots \sigma i_d) X(\sigma j_0 \cdots \sigma j_\lambda j_{\lambda+1} \cdots j_d),$$

where the sum ranges over the same permutations as above. Now, Theorem 1 says that $G_{d,n}$ can be expressed as the set of zeros of the particular such polynomials with $\lambda = d$. Consequently, $G_{d,n}$ can be expressed as the set of zeros of all the polynomials of the form (QP) as well. It can be shown formally (by a proof like that of (9) on page 379 of Vol. II of [**8**]) that each polynomial of the form (QP) is a linear combination with rational numbers as coefficients of the particular ones with $\lambda = d$ and it is an open question whether integers may be used as coefficients.

Let I be the ideal in the polynomial ring $R = k[\cdots, X(i_0 \cdots i_d), \cdots]$ generated by the polynomials of the form (QP) and let J be the subideal generated by the particular ones with $\lambda = d$. It can be shown that I is a prime ideal.

It then follows from the fact that I and J have the same zeros, that I is the radical of J. An interesting open question is whether I is always equal to J. They are equal in characteristic zero and would always be equal if the integers could be used as coefficients above.

The ring R/I is called the **homogeneous coordinate ring** of $G_{d,n}$ and plays an important role in the study of its geometry. The ring is naturally graded and the m-th graded piece consists of the residue classes of the homogeneous polynomials of degree m. Hodge and Littlewood (see [**8**], vol. II, chap. XIV, §9) have proved an explicit formula, known as the **postulation formula**, which expresses the dimension of m-th graded piece, for every m, as the value of a certain polynomial.

Igusa [**9**] (Theorem 1, p. 310) proved that R/I is a normal domain and derived several important results in invariant theory from this fact. The ring R/I is in fact a

unique factorization domain (see Samuel [**21**], Proposition 8.5, p. 38); this fact easily yields Severi's result that every $[(d+1)(n-d)-1]$-dimensional irreducible subvariety of $G_{d,n}$ is the intersection of $G_{d,n}$ and the set of zeros of a single homogeneous polynomial.

More recently, it has been proved (see Hochster [**6**] and Laksov [**16**]) that R/I is a Cohen-Macaulay ring. It follows by general principles that there is an exact sequence

$$O \to F_r \to F_{r-1} \to \cdots \to F_1 \to R \to R/I \to 0,$$

where the F_i are free R-modules and r is equal to $[N-(d+1)(n-d)]$. It is an interesting open problem to give an explicit natural such sequence, or in other words to find the syzygies of the ideal I of R.

Concerning section three.—For each Schubert subvariety $\Omega(A_0 \cdots A_d)$ of $G_{d,n}$, let $I(A_0 \cdots A_d)$ be the ideal of R generated by the quadratic polynomials of the form (QP) and the linear polynomials corresponding to the linear equations of Corollary 5. An important method for proving a result about the ring R/I is to prove more generally a corresponding result for each ring $R/I(A_0 \cdots A_d)$ by induction on the dimension of $\Omega(A_0 \cdots A_d)$. For example, this method is used to establish the postulation formula and the Cohen-Macaulay nature of R/I.

Another reason for interest in the rings $R/I(A_0 \cdots A_d)$ is that locally each $\Omega(A_0 \cdots A_d)$ can be described as the zeros of certain minors in the affine space of $(d+1) \times (n-d)$-matrices. For example, suppose that A_i consists of the points in $\boldsymbol{P}^n$ of the form $(p(0), \cdots, p(a_i), 0, \cdots, 0)$ and that for some $s \leqq d$ we have $a_i = (d-s+i)$ for $i = 0, \cdots, s$. Then Proposition 3 asserts that a point $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $G_{d,n}$ lies in $\Omega(A_0 \cdots A_d)$ if and only if $p(j_0 \cdots j_d)$ is zero whenever we have $a_i < j_i$ for some i At the end of section two, we noted that the points $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $G_{d,n}$ with $p(0 \cdots d) \neq 0$ are in natural bijective correspondence with the space of $(d+1) \times (n+1)$ matrices $[p_\alpha(j)]$ such that the $(d+1) \times (d+1)$ submatrix consisting of the first $(d+1)$ columns is the identity. Now, suppose that $p(j_0 \cdots j_d)$ is zero whenever we have $a_i < j_i$ for some i. Fixing $i \geqq s$ and considering all sequences

$$0 \leqq j_0 < \cdots < j_{i-1} \leqq d \leqq a_i < j_i < \cdots < j_d \leqq n$$

we easily conclude that all $(d-i+1) \times (d-i+1)$-minors of the $(d+1) \times (n-a_i)$-submatrix of $[p_\alpha(j)]$ consisting of the last $(n-a_i)$ columns are zero. Conversely, suppose that all such minors are zero whenever we have $i \geqq s$. Consider a determinant $p(j_0 \cdots j_d) = \det[p_\alpha(j_\beta)]$ with $a_i < j_i$ for some i. Since $i < s$ clearly implies $a_s < j_s$, we may assume $i \geqq s$. Then (Laplace) expansion of the determinant along the last $(d-i+1)$ columns shows that it is zero. Thus the points $(\cdots, p(j_0 \cdots j_d), \cdots)$ of $\Omega(A_0 \cdots A_d)$ with $p(0 \cdots d) \neq 0$ can be described as the zeros of all the $(d-i+1) \times (d-i+1)$-minors from the last $(n-a_i)$ columns for all $i \geqq s$ in the affine space of $(d+1) \times (n-d)$-matrices.

The zeros of determinantal equations are called **determinantal varieties** and have been studied for a long time (see Room [**19**]). Many of their properties can be easily deduced from corresponding properties of Schubert varieties. For example, let I' be the ideal of $k[X_{ij}]$ generated by the corresponding determinantal polynomials. The ring $k[X_{ij}]/I'$, known as the **coordinate ring** of the determinantal variety, is Cohen-Macaulay because a corresponding ring $R/I(A_0 \cdots A_d)$ is. Particular cases of this result were proved by Macaulay [**17**]; however, the general result was first established by Hochster and Eagon [**7**] without reference to the Grassmann manifold.

The syzygies of the ideal I' would be known if the syzygies of the corresponding ideal $I(A_0 \cdots A_d)$ were known, but in both cases it is an open problem to find the syzygies. In special cases they have been determined by Macaulay and Eagon—Northcott [**3**]. Recently, Kempf [**10**] has found a powerful way of determining syzygies, which gives an elegant treatment of some of the known cases and leads to the solution of new cases; (recently this was proved by Svanes [**24**]).

The Schubert varieties are, in general, singular. (Over the complex numbers, a singularity is a point where a subvariety is not a complex submanifold.) In fact, a point of $\Omega = \Omega(A_0 \cdots A_d)$ is singular if and only if the corresponding d-plane L in $\boldsymbol{P}^n$ satisfies $\dim(A_i \cap L) \geqq i$ for all i, as usual, and also $\dim(A_j \cap L) \geqq j+1$ for some j. Hence, the singular locus of Ω is a union of other Schubert varieties, and so the stratification $\Omega = \Omega_0 \supset \Omega_1 \supset \cdots \supset \Omega_m = \phi$, where Ω_i is the singular locus of Ω_{i-1}, is exceedingly well-behaved. Moreover, as we noted above, Ω is locally Cohen-Macaulay and so, since its singular locus is sufficiently small (of codimension at least two), Ω is also normal. Thus, the singularities are very nice. However, it remains to be proved that (trivial exceptions aside) these singularities are rigid—that any infinitesmal family varying an open piece of Ω must be analytically isomorphic to the trivial or product family. The rigidity is known in a very special case and it has applications to the theory of smoothing singularities (see Kleiman-Landolfi [**14**]).

Concerning section four.—Let us work over the complex numbers for a while. Most of the results of cohomology theory we used have become standard algebraic topology, but the assignment of a cohomology class to an algebraic subvariety of an algebraic manifold has not become standard. While early triangulations of such subvarieties have more recently been found unsatisfactory, today it is relatively easy to define the cohomology class either by using integration or relative (or local) cohomology and the difficulty lies in establishing the desired properties. A recent account of the theory is found in the article [**2**] of Borel and Haefliger.

The basis theorem was first proved by Ehresmann (see [**4**] §10, pp. 416–418). He observed that the Schubert varieties furnish a cellular decomposition of the Grassmann manifold because each Schubert variety contains an open subset which is an affine space (as we noted on the way to reformulating the basis theorem) and because the complement of this open set in the Schubert variety is the union of certain smaller Schubert varieties. The basis theorem then follows from some general results

about cell complexes which were included for this purpose and which have become standard. Ehresmann (see [4] §11, pp. 418–422) also proved the proposition complementing the basis theorem by a simple direct computation involving suitably chosen Schubert varieties to represent the Schubert cycles in question. He did not mention either the determinantal formula or Pieri's formula.

Another approach to Schubert calculus is by way of algebraic groups. When proving Proposition 4 in section three, we saw that the group $GL(n+1)$ of invertible $(n+1)\times(n+1)$-matrices acts on the Grassmann manifold $G_{d,n}$. It is easy to see that the action is transitive and that the d-plane in $\boldsymbol{P}^n$ whose points are of the form $(p(0),\cdots,p(d),0,\cdots,0)$ is left fixed by the matrices of the form

$$d+1\left\{\left[\begin{array}{c|c} \overbrace{\quad * \quad}^{d+1} & 0 \\ \hline * & * \end{array}\right]\right.$$

These matrices form a (parabolic) subgroup of $GL(n+1)$ and $G_{d,n}$ can obviously be considered as the quotient of $GL(n+1)$ by this subgroup. This observation suggests looking more generally at any quotient of a semi-simple algebraic group by a parabolic subgroup. The decomposition into Schubert cells can be correspondingly generalized by means of the Bruhat decomposition (see Borel [1], Theorem, page 347), and Kostant [15] has discovered a close connection between the (generalized) Schubert calculus and representation theory. In the case of the Grassmann manifold, the explicit formulas of (ordinary) Schubert calculus result from classical formulas of representation theory. In the general case, the situation is not fully understood.

Over an algebraically closed field of any characteristic, there are several purely algebraic theories which can take the place of classical cohomology. By far the most difficult to develop are the so called "Weil cohomologies" such as ℓ-adic cohomology. Over the complex numbers these theories are equivalent to classical cohomology and in any characteristic they have properties like the Künneth formula, Poincaré duality and classes for subvarieties. There are several less sophisticated theories (see Samuel [20]) which formally resemble the part of cohomology generated by the classes of subvarieties, but which may be weaker, that is, contain more information. The most popular of these is the weakest and is known as the **Chow ring** (see [23] and [25]). These theories constitute the topological and algebraic intersection theories (mentioned in section one), and we shall refer to any one of them as a **generalized cohomology theory.** At any rate, they are all equivalent for the Grassmann manifolds and the other varieties with cellular decompositions.

In Hodge-Pedoe [8], a generalized cohomology theory is developed in characteristic zero and the basis theorem for the Grassmann manifold $G_{d,n}$ is proved by induction on n. Then, the proposition complementing the basis theorem is proved by the same direct computation Ehresmann used. Next, Pieri's formula is deduced from the basis theorem and the proposition by another direct computation of the

same type. Finally, the determinantal formula is deduced formally from Pieri's formula. In fact, with a generalized cohomology theory and the basis theorem given, the remaining three results can always be derived without difficulty in this way in any characteristic.

The Grassmann manifold $G_{d,n}$ can obviously be thought of as representing the $(d+1)$-dimensional (vector) subspaces of an $(n+1)$-dimensional vector space. From this point of view, it is natural to consider the trivial vector bundle of rank $(n+1)$ on $G_{d,n}$ and its canonical subbundle E whose fiber over a point of $G_{d,n}$ is the $(d+1)$-dimensional (vector) subspace of the $(n+1)$-dimensional vector space represented by the point. This subbundle E is universal in the sense that for any variety X and for any subbundle of rank $(d+1)$ of the trivial bundle of rank $(n+1)$ on X, there is a unique map of X into $G_{d,n}$ such that the subbundle E on $G_{d,n}$ induces the given subbundle on X.

A general theory of Chern classes with values in any generalized cohomology theory has been worked out (see Grothendieck [**5**]), and the special Schubert cycle $\sigma(h)$ is exactly the $(n-d-h)$-th Chern class of the quotient of the trivial bundle of rank $(n+1)$ on $G_{d,n}$ by the universal subbundle (see Kleiman [**12**], p. 297). The results of Schubert calculus now yield a description of the generalized cohomology of $G_{d,n}$ as the ring generated by these Chern classes. Grothendieck (see [**23**], Théorème 1, p. 4–19) has given a formal derivation of this description, without any mention of Schubert varieties or cycles.

The determinantal formula is related to a very useful formula of Porteous in differential geometry and it appears in the study of the singularities of a map (see [**18**]). The determinantal formula is also the key to proving the existence of certain special divisors on curves (see Kempf [**11**] and Kleiman-Laksov [**13**]), and in his article [**11**], Kempf gives a nice direct proof of the formula.

Another source of interest in Schubert varieties is the problem of smoothing cycles. The problem is to show that the class of any subvariety Z of a nonsingular algebraic variety V is the difference of two classes each the class of a nonsingular subvariety. When $\dim(Z) < (\dim(V)+2)/2$ holds, then some multiple of the class of Z is such a difference and the proof involves a careful study of the geometry of certain Schubert varieties (see Kleiman [**12**]). However, it is suspected that the general problem has a negative solution and in fact that the Schubert cycle $\sigma(1)$ on the Grassmann manifold $G_{2,5}$ is not the difference of two cycles each the class of a nonsingular subvariety, nor is any multiple of $\sigma(1)$.

The examples from enumerative geometry we considered, while simple, illustrate fairly well the use of Schubert calculus. Classically relatively complicated geometric situations were studied. They often involved tangency conditions such as requiring a line to be an n-fold tangent to a given curve or requiring a line to intersect a given surface and lie in the tangent plane of the surface at the point of intersection. In principle, the method is always the same: describe the problem in terms of subvarieties of a Grassmann manifold; find the degrees of each subvariety; and use the

formulas of Schubert calculus to compute the product of the classes of the subvarieties. Moreover, each degree is the number of points of intersection of a subvariety with a certain Schubert variety, or in other words, it is the number of solutions to a certain simpler enumerative problem. In practice, finding the degrees can be difficult and may, as in the case of tangency conditions, involve more sophisticated algebraic geometry.

Although we have given the "principle of conservation of number" a rigorous mathematical interpretation, it is usually difficult to use it because it is difficult to know what the correct multiplicities are. For example, consider the lines in $\boldsymbol{P}^3$ intersecting lines L_1, L_2, L_3, L_4; how can we tell by direct geometric means that if L_1, L_2 and L_3 are skew and L_4 intersects each of them, then the one solution (found at the end of section three) should be counted with multiplicity two, or, for that matter, how can we tell that if L_1 intersects L_2 and L_3 intersects L_4, then the two solutions (found in section one) should each be counted with multiplicity one? In the general case of an enumerative problem, it is possible to prove, in characteristic zero and often in characteristic p, that the solutions all appear with multiplicity one. Thus, for example, we may assert that the number of distinct lines in $\boldsymbol{P}^3$ meeting four curves C_1, C_2, C_3, C_4 of degree $\delta_1, \delta_2, \delta_3, \delta_4$ is, in general, $2\delta_1\delta_2\delta_3\delta_4$ and that two quadrics in $\boldsymbol{P}^4$ have, in general, sixteen lines in common. In analyzing the latter example, we used geometric means to see that there are four lines which simultaneously lie on a general quadric, lie on a general 3-plane and intersect a general line in this 3-plane. Here we are able to say that each solution appears with multiplicity one because the quadric, the 3-plane and the line in the 3-plane satisfy no special conditions.

In more abstract terms, we can assert in characteristic zero (see [**8**], p. 338) that for any two irreducible subvarieties X and Y of $G_{d,n}$, the components all appear with multiplicity one in the intersection of X and the image of Y under the linear transformation of $G_{d,n}$ into itself induced by any sufficiently general invertible $(d+1) \times (n+1)$-matrix. It would be interesting to know what happens in characteristic p.

Supported in part of NSF GP 28936.

References

1. A. Borel, Linear Algebraic Groups, Benjamin, New York, 1969.

2. A. Borel and A. Haefliger, La classe d'homologie fondamentale d'un espace analytique, Bull. Soc. Math. France, 89 (1961).

3. J.A. Eagon and D.G. Northcott, Ideals defined by matrices and a certain complex associated with them, Proc. Roy. Soc. A, 269 (1962).

4. C. Ehresmann, Sur la topologie de certains espaces homogènes, Ann. Math., 35 (1934).

5. A. Grothendieck, La théorie des classes de Chern, Bull. Soc. Math. France, 86 (1958).

6. M. Hochster, Grassmannians and their Schubert subvarieties are arithmetically Cohen-Macaulay, (to appear).

7. M. Hochster and J. A. Eagon, Cohen-Macaulay rings, invariant theory, and the generic perfection of determinantal loci, (to appear).

8. W. V. D. Hodge and D. Pedoe, Methods of algebraic geometry, vol. I and II, Cambridge University Press, 1953.

9. J. I. Igusa, On the arithmetic normality of the Grassmann variety, Proc. Nat. Acad. Sci., 40 (1954).

10. G. Kempf, The singularities of certain varieties in the Jacobian of a curve, (to appear).

11. G. Kempf, Schubert methods with an application to algebraic curves, Stichting mathematisch Centrum, Amsterdam, 1971.

12. S. L. Kleiman, Geometry of Grassmannians and applications..., Publ. Math. I. H. E. S. No. 36, Paris (1969).

13. S. L. Kleiman and D. Laksov, On the existence of special divisors, J. Amer. Math. Soc. (to appear).

14. S. L. Kleiman and J. Landolfi, Geometry and deformation of special Schubert varieties, Compositio Math. (to appear).

15. B. Kostant, Lie algebra cohomology and the generalized Borel-Weil theorem, Ann. Math., no. 2, 74 (1961) Lie algebra cohomology and generalized Schubert cells, Ann. Math, no. 1, 77 (1963).

16. D. Laksov, Concerning the arithmetic Cohen-Macaulay character of Schubert schemes, Acta Math., (to appear).

17. F. S. Macaulay, The algebraic theory of modular systems, Cambridge Tracts, 19 (1916).

18. *Proceedings of Liverpool Singularities-Symposium* 1, Vol. 192. Lecture notes in mathematics, Springer Verlag, New York, 1971.

19. T. G. Room, The geometry of determinental loci, Cambridge Univ. Press, 1938.

20. P. Samuel, Relations d'équivalence en géométrie algébrique, Proc. Intern. Congress of Math., Cambridge Univ. Press, 1960.

21. ———, Lectures on unique factorization domains, Tata Inst. of Fund. Research, Bombay, 1964.

22. H. Schubert, Kalkül der Abzählenden Geometrie, Teubner, Leipzig, 1874.

23. Seminaire C. Chevalley, Anneaux de Chow, Paris, 1958.

24. T. Svanes, Thesis, M. I. T. 1972.

25. *Théorie des intersections et théorème de Riemann-Roch.* S. G. A. 6 (1966/67), Lecture notes in mathematics, Vol 225 Springer-Verlag, New York, 1971.

PRIME FACTORS OF CONSECUTIVE INTEGERS

E. F. ECKLUND, JR., Northern Illinois University,
and R. B. EGGLETON, The University of Calgary

For each positive integer k, there exists a corresponding positive integer m such that in any sequence of m consecutive integers greater than k there is at least one having a prime factor greater than k. A simple demonstration of this fact comes from a modification of Euclid's proof that there are infinitely many primes. Let P be the product of all primes not larger than k, and let $a_1 < a_2 < \cdots < a_{\phi(P)}$ be the positive integers not greater than P which are prime relative to P. For any a_i and any integer r, the number $rP + a_i$ is prime relative to P so, if greater than 1, has

all its prime factors greater than k. The greatest gap which occurs between such numbers is clearly finite, so any larger number could be chosen as m. If $f(k)$ denotes the smallest possible choice of m, and we define $a_{\phi(P)+1} = P + a_1$, then

$$f(k) \leqq \max\{a_{i+1} - a_i : 1 \leqq i \leqq \phi(P)\}.$$

The elegant result $f(k) \leqq k$ is the content of a theorem proved independently by Sylvester [**11**] and Schur [**9**]. In providing an elementary proof, Erdös [**1**] stated the Sylvester-Schur Theorem in the following form: *For positive integers n and k with $n \geqq 2k$, $\binom{n}{k}$ has a prime factor greater than k.* This follows from a stronger theorem, due essentially to Erdös [**3**]:

THEOREM. *If $k \geqq 202$ and $n \geqq 2k$, then $\binom{n}{k} > n^{\pi(k)}$, where $\pi(k)$ denotes the number of primes less than or equal to k.*

Before proving this theorem, we note the following bound:

LEMMA. $\displaystyle\binom{2k}{k} > \frac{4^k}{k}$ *for* $k \geqq 4$.

Proof. If $k = 4$, $\binom{2k}{k} = \binom{8}{4} = 70 > 64 = 4^k/k$. Assume $\binom{2k}{k} > 4^k/k$; then

$$\binom{2k+2}{k+1} = \frac{(2k+2)(2k+1)}{(k+1)^2}\binom{2k}{k} > \frac{2(2k+1)}{k+1}\cdot\frac{4^k}{k}$$

$$> \frac{4k\cdot 4^k}{k(k+1)} = \frac{4^{k+1}}{k+1}.$$

Hence the lemma follows by induction. ∎

Proof of the theorem: First we establish that $\binom{2k}{k} > (2k)^{\pi(k)}$, using the bound $k/(\log k - {}^3/_2) > \pi(k)$ for $k > e^{3/2}$, due to Rosser and Schoenfeld [**8**]. Suppose

$$\frac{4^k}{k} > (2k)^{k/(\log k - 3/2)}.$$

Taking logarithms, this supposition is equivalent to

$$k\log 4 - \log k > (\log 2 + \log k)k/(\log k - \tfrac{3}{2}),$$

which is true for $k \geqq 1414$. Therefore, using the lemma,

$$\binom{2k}{k} > \frac{4^k}{k} > (2k)^{k/(\log k - 3/2)} > (2k)^{\pi(k)} \text{ for } k \geqq 1414.$$

Checking by computer, we have verified that $\binom{2k}{k} > (2k)^{\pi(k)}$ for $202 \leqq k \leqq 1413$. Hence $\binom{2k}{k} > (2k)^{\pi(k)}$ for $k \geqq 202$.

Now assume for fixed k that n satisfies $\binom{n}{k} > n^{\pi(k)}$. For any $s \geqq 1$,

$$k > \pi(k) \geqq 1 + \frac{\pi(k)-1}{s} = \frac{\pi(k)+s-1}{s},$$

whence the product over $1 \leqq s \leqq r$ yields $k^r > \binom{\pi(k)+r-1}{r}$. Then

$$\binom{n+1}{k} = \left(1 - \frac{k}{n+1}\right)^{-1}\binom{n}{k} > \left(1 - \frac{k}{n+1}\right)^{-1} n^{\pi(k)}$$

$$= n^{\pi(k)} \sum_{r \geqq 0} \frac{k^r}{(n+1)^r} > n^{\pi(k)} \sum_{r \geqq 0} \binom{\pi(k)+r-1}{r} \frac{1}{(n+1)^r}$$

$$= n^{\pi(k)}\left(1 - \frac{1}{n+1}\right)^{-\pi(k)} = (n+1)^{\pi(k)}.$$

The theorem follows by induction on n. ∎

COROLLARY. *For each* $k \geqq 1$ *there is an* n_k *such that* $\binom{n}{k} > n^{\pi(k)}$ *just if* $n > n_k$. *For* $k \geqq 202$, $n_k < 2k$.

The Sylvester-Schur theorem follows from this result in virtue of the following lemma [**1**]:

LEMMA. *If* $\binom{n}{k}$ *is divisible by a prime-power* p^α, *then* $p^\alpha \leqq n$.

Proof. In $\binom{n}{k}$ the prime p has exponent

$$\alpha = \sum_{1 \leqq s \leqq r} \left(\left[\frac{n}{p^s}\right] - \left[\frac{k}{p^s}\right] - \left[\frac{n-k}{p^s}\right]\right),$$

where $p^r \leqq n < p^{r+1}$; and each of the r summands has value 0 or 1. ∎

Proof of the Sylvester-Schur Theorem: This lemma shows the contribution to $\binom{n}{k}$ from primes not exceeding k is at most $n^{\pi(k)}$, so $\binom{n}{k}$ has a prime factor greater than k if $n > n_k$. The results already established imply the Sylvester-Schur Theorem in all cases except for the 638 pairs (n,k) with $n \geqq 2k$ such that $\binom{n}{k} \leqq n^{\pi(k)}$. These pairs can be deduced from Table 1, which lists those pairs (n_k, k) for which $n_k \geqq 2k$. A simple check for these exceptional pairs completes the proof. ∎

Stronger bounds on $f(k)$ have been established using analytic methods. Erdös [**2**] showed that there exists a constant $c_1 > 1$ such that

$$f(k) \leqq c_1 \frac{k}{\log k}.$$

The best upper bound obtained thus far is

$$f(k) \leqq (\tfrac{1}{2} + \varepsilon) \frac{k}{\log k}$$

for any $\varepsilon > 0$ and $k > k_\varepsilon$, due to the work of Ramachandra [**6**] and Tijdeman [**12**].

TABLE 1. Pairs (n_k, k) with $n_k \geqq 2k$.

k	3	5	6	7	8	9	11	12	13	14	15—16	17	18	19
n_k	8	15	13	23	20	19	27	26	35	34	33	41	40	50

k	20	21—22	23	24	25—27	29—30	31	32—35	37—39	41—42	43—45
n_k	49	48	57	56	55	63	72	71	79	87	96

k	46	47	48—52	53—54	55—56	59—60	61—65	67—69	71—72	73
n_k	95	105	104	113	112	121	130	139	148	158

k	74–78	79–82	83	84–87	89	90–92	97	101	103–106
n_k	157	166	176	175	185	184	194	203	213

k	107–108	109	110–112	113–117	118–120	139–140	199–201
n_k	222	233	232	242	241	280	402

A lower bound is also known; by a result of Rankin [7], it follows that for some constant $c_2 > 0$,

$$f(k) > c_2 \frac{\log k (\log\log k)\,(\log\log\log\log k)}{(\log\log\log k)^2}.$$

Erdös suggested that the asymptotic behavior of $f(k)$ may be $f(k) \sim (\log k)^2$ for $k \to \infty$. This is based on the conjectured size of the gap between successive primes, as $f(k)$ seems not to be significantly larger than the difference between the first two primes greater than k.

The actual evaluation of $f(k)$ for $k \leqq 10$ was done by Utz [**13**], and extended to $k \leqq 42$ by Lehmer [**5**]. Table 2 summarizes their results.

TABLE 2. Values of $f(k)$

k	1	2	3–4	5–12	13–40	41–42
$f(k)$	1	2	3	4	6	7

Questions raised by the study of $f(k)$ are given a slightly wider setting if we introduce $g(k,m)$, defined to be the smallest nonnegative integer with the property that in every sequence of m consecutive integers greater than $g(k,m)$ there is at least one with a prime factor greater than k. Thus the Sylvester-Schur Theorem takes the form $g(k,k) \leqq k$, and we recover $f(k)$ via

$$f(k) = \min\{m : g(k,m) \leqq k\}.$$

(Our $g(k,m)$ is actually a proper extension of a certain $U_m(n)$ used by Lehmer in **[5]**. If p_n denotes the nth prime, the connection is simply $U_m(n) = g(p_n, m)$ provided $1 < m < f(p_n)$, the condition for $U_m(n)$ to exist.)

Evidently there is no integer $g(k,1)$ if $k \geqq 2$. However, it is a remarkable fact that for any given k there are only finitely many pairs of consecutive integers not divisible by any prime greater than k: this follows from a theorem of Størmer **[10]**. Thus $g(k,2)$ exists for every k, and *a fortiori*, so does $g(k,m)$ for $m \geqq 2$. Indeed Lehmer's results in **[4]** imply that, for some constant $c_3 > 0$,

$$\log g(k,2) < c_3 k^2 \exp(k/2).$$

Adapting a construction of Erdös **[3]**, the lower bound $g(k,2) > k^2$ can be established when $k \geqq 5$. For if $k+1$ and $k+2$ are composite, then $(k+1)^2 - 1 = k(k+2)$ and $(k+1)^2$ are consecutive numbers with no prime factor greater than k. Now suppose $k \geqq 5$. If $k+1$ is prime, k is even and at least one of $k+3$ and $k+5$ is composite, so either $(k+2)(k+4)$ and $(k+3)^2$ are consecutive numbers with no prime factor greater than k, or else $(k+4)(k+6)$ and $(k+5)^2$ are such a pair. Similar reasoning applies if $k+2$ is prime. Indeed, a similar argument shows $g(k,2) > k^4$ for infinitely many k. Erdös has suggested that $g(k,2) \sim \exp(\sqrt{k})$ may be true.

More generally, although the asymptotic behaviour of $g(k,m)$ remains unknown, an argument of Erdös **[3]** generalizes to give a lower bound on this function. It uses the following well-known result for the sum of reciprocals of primes smaller than n:

$$\sum_{p<n} \frac{1}{p} - \log\log n \to c_4,$$

where c_4 is a small positive constant. Let n, n' be integers satisfying $n > n' > g(k,m)$. If N denotes the number of integers r satisfying $n' \leqq r < n$ which are divisible by at least one prime greater than k, the definition of $g(k,m)$ ensures

$$\left[\frac{n-n'}{m}\right] \leqq N \leqq \sum_{k<p<n}\left(\left[\frac{n-1}{p}\right] - \left[\frac{n'-1}{p}\right]\right).$$

Given $\varepsilon > 0$, it follows for all sufficiently large k that

$$\frac{1}{m} - \varepsilon < \log\log n - \log\log k,$$

and therefore

$$\log n > \exp\left(\frac{1}{m} - \varepsilon\right)\log k.$$

Choosing n comparable with $g(k,m)$, this implies the desired bound: for $\varepsilon > 0$ and $k > k_\varepsilon$,

$$\log g(k,m) > \exp\left(\frac{1}{m} - \varepsilon\right) \log k.$$

It is clear that $g(k+1,m) \geqq g(k,m)$ for every k. In fact, $g(k,m) = g(p,m)$ where p is the largest prime not greater than k. Also $g(1,1) = 0$, so it is sufficient to specify values of $g(k,m)$ for prime values of k. Table 3, which is based on Lehmer's results in [**5**], has been constructed accordingly. (However, we note that $g(11,3)=98$, correcting the value 54 given in [**5**].) Since $g(k,f(k)) \leqq k$, it is easy to see $g(k,f(k)+r) = \max\{0, g(k,f(k)) - r\}$ for every $r \geqq 0$; Table 3 has been abbreviated by omitting values of $g(k,m)$ for $m > f(k)$.

TABLE 3. Values $g(k, m)$

k \ m	2	3	4	5	6	7
2	1					
3	8	2				
5	80	8	3			
7	4 374	48	7			
11	9 800	98	9			
13	123 200	350	63	24	11	
17	336 140	440	63	48	13	
19	11 859 210	2 430	168	48	17	
23	11 859 210	2 430	322	48	23	
29	177 182 720	13 310	322	54	25	
31	1 611 308 699	13 454	1 518	152	31	
37	3 463 199 999	17 575	1 518	152	35	
41	63 927 525 375	212 380	1 680	286	285	36

It was noted by J. L. Selfridge that if $a-1$, a, $a+1$ are consecutive integers with no prime factors greater than k, then a^2-1, a^2 are consecutive integers with the same property. Hence we can deduce that

$$g(k,2) \geqq g(k,3)(g(k,3)+2).$$

(In particular, $g(11,2) = 9800$ implies $g(11,3) \leqq 98$.) It is also simple to observe that any sequence of m consecutive integers with no prime factor greater than p_n, the nth prime, cannot contain more than one multiple of p_n, since $g(p_n, p_n) \leqq p_n$ and $p_{n+1} < 2p_n$. Therefore such a sequence must contain at least $[m/2]$ consecutive integers with no prime factor greater than p_{n-1}, so

$$g(p_n, m) \leqq g(p_{n-1}, [m/2]).$$

In particular, $g(43,6) \leqq g(41,3) = 212\,380$. Using this bound, we obtained $g(43,6) = 340$ and $g(43,7) = 40$ by a simple computer search. Thus $f(k) = 7$ for $43 \leqq k \leqq 46$.

For any integer $r \geqq 2$, let $A_r(k)$ be the first occurring sequence of maximal length among all sequences of consecutive integers which have no prime factor greater than k, and smallest member a satisfying $k < a \leqq rk$. Let $a_r(k)$ and $|A_r(k)|$ denote the smallest integer and number of integers in $A_r(k)$, respectively. Table 4 specifies $a_r(k)$ and $|A_r(k)|$ for all primes $p \leqq 283$, with $r = \min\{100,[13000/p]\}$. For the larger primes a computer was used. Clearly $f(p) \geqq |A_r(p)| + 1$, with equality for all sufficiently large r. In fact, Table 4 shows equality holds for all primes $p \leqq 37$ when $r \geqq 2$, except for $p = 3$ (when $r \geqq 3$ applies). Equality for $p = 41$ or 43 holds when $r \geqq 7$. It seems likely that equality has been achieved in Table 4 for $47 \leqq p \leqq 283$.

It is not known if $f(k)$ is monotone. Evidence that it is not is provided by Table 4. which suggests that $f(113) = 14$ and $f(127) = 12$. If p is the largest prime not greater than k, we have noted that $g(k,m) = g(p,m)$. With $f(k)$, the corresponding situation is more complicated. Correcting a remark in [**13**], we note that $f(k) \leqq f(p)$, and equality holds only if $g(p,f(p)-1) > k$; it is not at all clear that this inequality always holds. (Table 4 suggests that $114, 115, \cdots, 126$ is the last sequence of 13 consecutive integers none of which has a prime factor greater than 113, and if so, $f(113) = 14$ and $f(114) = 13$.)

We close with an elementary proof of an interesting bound on g. Let $(n)_m$ denote the falling factorial $n(n-1)\cdots(n-m+1)$; we write $p^\alpha \| n$ to mean $p^\alpha | n$ and $p^{\alpha+1} \nmid n$.

THEOREM. $g(p_n, n+1) \leqq (n)_{\pi(n)}$.

Proof. Suppose T is a set of $n+1$ consecutive integers such that none is a multiple of any prime greater than p_n. Let S be a maximal subset of T with the property that for each $s \in S$ there is a corresponding prime p such that $p^\alpha \| s$, no other member of S is a multiple of p^α, and no member of T is a multiple of $p^{\alpha+1}$. Now $|S| \leqq n$, since S has no more elements than there are primes less than or equal to p_n. Thus we can select some $a \in T \setminus S$. Since T can contain at most one multiple of any prime greater than n, the prime factors of a do not exceed n. With $1 \leqq i \leqq \pi(n)$, let $p_i^{\alpha_i} \| a$; correspondingly there exist $b_i \in T \setminus \{a\}$ such that $p_i^{\beta_i} \| b_i$ and $\beta_i \geqq \alpha_i$. Let $I_1, I_2, \cdots, I_m$ be a partition of the set $\{1,2,\cdots,\pi(n)\}$ such that $b_i = b_j$ if and only if $\{i,j\} \subseteq I_r$ for some r. Since $p_i^{\alpha_i} | p_i^{\beta_i}$, for $1 \leqq r \leqq m$ we have

$$\Big(\prod_{i \in I_r} p_i^{\alpha_i}\Big) \Big| (a - b^{(r)})$$

where $b^{(r)} = b_i$ for each $i \in I_r$. Hence

$$a = \prod_{1 \leqq i \leqq \pi(n)} p_i^{\alpha_i} = \prod_{1 \leqq r \leqq m} \prod_{i \in I_r} p_i^{\alpha_i} \leqq \prod_{1 \leqq r \leqq m} |a - b^{(r)}|.$$

Now $m \leqq \pi(n)$, and induction on m easily shows the last product cannot exceed

the falling factorial $(n)_{\pi(n)}$. Hence T necessarily contains a member not greater than $(n)_{\pi(n)}$, and the theorem follows. ∎

TABLE 4. Specification of $A_r(p)$.

p	2	3	5–7	11	13–23	29–31
$a_r(p)$	4	8	8	14	24	32
$\lvert A_r(p)\rvert$	1	2	3	3	5	5

p	37	41—43	47—53	59	61–113
$a_r(p)$	48	285	90	114	114
$\lvert A_r(p)\rvert$	5	6	7	8	13

p	127–149	151–211	223–263	269–283
$a_r(p)$	200	294	1330	524
$\lvert A_r(p)\rvert$	11	13	15	17

This paper was substantially completed while the authors were participants in the 1971 N.S.F. Advanced Science Seminar on Combinatorial Theory at Bowdoin College. We express thanks to Bowdoin College for making computing facilities available to us. We are also indebted to Professor P. Erdös for several valuable comments, and especially for the new proof of the Sylvester-Schur Theorem, the main ideas in the argument being due to him.

References

1. P. Erdös, A theorem of Sylvester and Schur, J. London Math. Soc., 9(1934) 282–288.

2. ———, On consecutive integers, Nieuw. Arch. Wisk., 3(1955) 124–128.

3. ———, Private communication.

4. D. H. Lehmer, On a problem of Størmer, Illinois J. Math., 8(1964) 57–79.

5. ———, The prime factors of consecutive integers, this MONTHLY, 72(1965) no. 2, part II, 19–20.

6. K. Ramachandra, On numbers with a large prime factor III, Acta Arith., 19(1971) 49–62.

7. R. A. Rankin, The difference between consecutive prime numbers, J. London Math. Soc., 13(1938) 242–247.

8. J. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois J. Math., 6(1962) 64–94.

9. I. Schur, Einige Sätze über Primzahlen mit Anwendungen auf Irreduzibilitätsfragen, S. B. Deutsch. Akad. Wiss. Berlin Kl. Math. Phys. Tech., 23(1929) 1–24.

10. G. Størmer, Quelques théorèmes sur l'équation de Pell $x^2 - Dy^2 = \pm 1$ et leurs applications, Videnskabs-Selskabets Skrifter, Christiania, 1897, no. 2, 48 pp.

11. J. J. Sylvester, On arithmetical series, Messenger Math., 21(1892) 1–19, 87–120; Collected Mathematical Papers, vol. 4, 1912, pp. 687–731.

12. R. Tijdeman, On the maximal distance of numbers with a large prime factor, to appear.

13. W. R. Utz, A conjecture of Erdös concerning consecutive integers, this MONTHLY, 68(1961) 896–897.

THE TANGENT BUNDLE OF A TOPOLOGICAL MANIFOLD

RICHARD LASHOF, University of Chicago

A survey by R. Schultz of recent results on topological manifolds has appeared in the Monthly [5]. One important tool for the study of such manifolds has been the generalization to topological manifolds of the notion of the tangent bundle of a smooth surface (i.e., manifold).

Given a smooth surface M^n contained in euclidean $n + k$ space R^{n+k}, the tangent bundle $T(M)$ is the family of vectors in R^{n+k} tangent to the surface at some point. Explicitly,

$$T(M) = \left\{(x_0, y_0) \in R^{n+k} \times R^{n+k} \,\middle|\, x_0 \in M, \text{ and } y_0 = \left.\frac{dx}{dt}\right|_0,\right.$$

where $x(t)$ is a smooth curve contained in M with $x(0) = x_0\}$.

Since this definition involves the derivative, there is an obvious difficulty in trying to define the tangent bundle of a topologically embedded surface. The removal of this difficulty is a good case study in the use of abstraction in the development of mathematics. First we must abstract the notion of surface so that it is independent of the particular embedding in euclidean space. This is the notion of manifold. Second, we must abstract the notion of bundle. Then we must relate the two notions so as to generalize the situation described above.

(For an extended bibliography see [5], we give here only some general reference books and a few references not in [5].)

1. Notion of an *n*-manifold [1]. Recall that an n-dimensional smooth surface in R^{n+k} is defined locally by a smooth map (i.e., C^∞ function) $h\colon U \to R^{n+k}$, where U is an open set in R^n, and the Jacobian $J(h)$ has rank n at each point of U.

DEFINITION 1.1: A (topological) **manifold** M^n is a Hausdorff, second countable space, such that each point of M^n has an open neighborhood homeomorphic to an open subset of R^n. This last is equivalent to saying M^n has an open cover $\{V_\alpha\}_{\alpha\in A}$, such that for each α there is a homeomorphism h_α of an open subset U_α of R^n onto V_α. Each pair (h_α, U_α) is called a **chart**, and the family $\{(h_\alpha, U_\alpha)\}_{\alpha\in A}$ is called an **atlas.**

Obviously, an embedded surface M^n is a manifold, where $h\colon U \to M \subset R^{n+k}$ defines a chart.

Now a smooth surface has the additional property that if $h_\alpha\colon U_\alpha \to R^{n+k}$ and $h_\beta\colon U_\beta \to R^{n+k}$ are two local defining smooth maps, then $h_\beta^{-1} h_\alpha$ is smooth on $h_\alpha^{-1} h_\beta(U_\beta) \subset R^n$ to $h_\beta^{-1} h_\alpha(U_\alpha) \subset R^n$.

Richard Lashof received his Columbia University Ph.D. under R. L. Taylor and R. Kadison. He has been at the University of Chicago since then where he is now a Professor and where he served several years as Department Chairman. He held a Senior NSF Fellowship at the Institute for Advanced Study and he spent a year at Oxford University. His main research is topology and differential geometry. *Editor.*

DEFINITION 1.2: A **smooth structure** on a topological manifold M^n is an atlas $\{(h_\alpha, U_\alpha)\}_{\alpha \in A}$ such that $h_\beta^{-1} h_\alpha$ is smooth, all $\alpha, \beta \in A$. Two smooth structures are called equivalent if their reunion defines a smooth structure. A manifold M^n together with an equivalence class of smooth structures is called a **smooth manifold.**

We can now define a **smooth function** f: $M^n \to R^k$ on a smooth manifold to be a function such that $f \circ h_\alpha$ is smooth for each chart (h_α, U_α) of a smooth structure for M. It is easy to see that this is independent of the smooth structure in the equivalence class. Similarly, a function f: $M^n \to N^k$ between smooth manifolds is called smooth if $k_\beta^{-1} f \circ h_\alpha$ is smooth for each chart (h_α, U_α) and (k_β, W_β) of smooth structures on M and N respectively. A smooth map f: $M^n \to N^n$ is called a **diffeomorphism** if f is one to one and f^{-1} is smooth.

2. Notion of an R^n-bundle, [6]. The condition mentioned above, that $J(h)$ has rank n, for a local defining map of a smooth surface; implies that for each $x \in M^n$ the set of vectors in R^{n+k} tangent to M at x forms an n-plane $T_x(M)$. In fact, if $u_0 \in U$, the vectors tangent to curves $u(t)$ through u_0 in R^n is simply R^n translated to u_0, and if $x(t) = h(u(t))$,

$$\left.\frac{dx}{dt}\right|_0 = J(h)_{u_0} \left.\frac{du}{dt}\right|_0 .$$

Thus $T(M) = \bigcup_{x \in M} T_x(M)$ is a collection of n-planes. Further, we have continuous maps $s: M \to T(M)$ and $p: T(M) \to M$, where $s(x)$ is the zero vector in $T_x(M)$ and $p| \, T_x(M) = x$. Thus $ps = id_M$.

DEFINITION 2.1: Given topological spaces X and E and maps $X \xrightarrow{s} E \xrightarrow{p} X$ such that $ps = id_X$, (s, p) is called an **R^n-bundle** over X if for each $x \in X$ there is an open neighborhood V and a homeomorphism k: $V \times R^n \to p^{-1}(V)$ such that the composite maps

$$V \times R^n \xrightarrow{k} p^{-1}(V) \xrightarrow{p} V, \text{ and } V \xrightarrow{s|V} p^{-1}(V) \xrightarrow{k^{-1}} V \times R^n$$

are respectively projection onto the first factor, and the injection $x \to (x, 0)$, $x \in V$.

(k, V) is called a **local trivialization,** and for each $x \in X$, $p^{-1}(x)$ is called the **fibre** over x. Note that $p^{-1}(x)$ is homeomorphic to R^n.

The tangent bundle of a smooth surface is an R^n-bundle where k: $V \times R^n \to p^{-1}(V)$ is given by

$$k(v, y) = (v, J(h)_{h^{-1}(v)} y)$$

for h: $U \to V \subset M \subset R^{n+k}$ a smooth local defining map. Since $J(h)$ is a linear map this suggests:

DEFINITION 2.2: A **vector bundle structure** on an R^n-bundle (s, p) is a family $\{(k_\alpha, V_\alpha)\}_{\alpha \in A}$ of local trivializations such that

(1) $\{V_\alpha\}_{\alpha \in A}$ is an open covering of X,

(2) $k_\beta^{-1}k_\alpha\colon V_\alpha \cap V_\beta \times R^n \to V_\alpha \cap V_\beta \times R^n$ satisfies $(k_\beta^{-1}k_\alpha)_x\colon R^n \to R^n$ is a linear isomorphism, when $(k_\beta^{-1}k_\alpha)_x$ is defined by $(k_\beta^{-1}k_\alpha)_x(y) = \mathrm{proj}_2 k_\beta^{-1}(k_\alpha(x,y))$.

Two vector bundle structures on (s,p) are called **equivalent** if their reunion is a vector bundle structure. An n-dimensional vector bundle is an R^n-bundle together with an equivalence class of vector bundle structures. Note that each fibre of a vector bundle has a well-defined linear structure.

If $X_1 \xrightarrow{s_1} E_1 \xrightarrow{p_1} X_1$ and $X_2 \xrightarrow{s_2} E_2 \xrightarrow{p_2} X_2$ are two R^n-bundles, a pair of maps (ϕ,f), $\phi\colon E_1 \to E_2$, $f\colon X_1 \to X_2$, is called an **R^n-bundle map** if:

(1) $p_2\phi = fp_1$,

(2) $s_2 \circ f = \phi s_1$,

(3) $\phi\,|\, p_1^{-1}(x_1);\ p_1^{-1}(x_1) \to p_2^{-1}(f(x_1))$ is a homeomorphism for all $x_1 \in X_1$.

Further, if (s_1,p_1) and (s_2,p_2) are vector bundles, a bundle map is called **linear** if it is linear on fibres. A (linear) bundle map is called a (vector) **bundle equivalence** if $X_1 = X_2 = X$, $f = id_X$. It is not difficult to show that if (ϕ, id) is a (vector) bundle equivalence, then ϕ is a homeomorphism and (ϕ^{-1}, id) is a (vector) bundle map.

If M^n is a manifold, and $M \xrightarrow{s} E \xrightarrow{p} M$ is a R^k-bundle over M, it follows from the local trivializations that E is an $n+k$ dimensional manifold.

If M^n is a smooth manifold, and (s,p) is a vector bundle over M, then it can be shown that E has a smooth structure, unique up to equivalence, such that the local trivializations are diffeomorphisms.

3. The tangent vector bundle of a smooth manifold [1]. Having abstracted the notions of smooth manifold and vector bundle, we define the **tangent vector bundle** of a smooth manifold M as follows: Let $\{(h_\alpha, U_\alpha)\}$ define a smooth structure for M. If $x \in h_\alpha(U_\alpha)$, the set of vectors tangent to curves through $h_\alpha^{-1}(x) \in R^n$ is isomorphic to R^n under translation of the origin to $h_\alpha^{-1}(x)$. If x also is in $h_\beta(U_\beta)$, then $h_\beta^{-1}h_\alpha$ sends smooth curves through $h_\alpha^{-1}(x)$ to smooth curves through $h_\beta^{-1}(x)$ and $J(h_\beta^{-1}h_\alpha)_{h^{-1}(x)}$ gives a linear isomorphism on tangent vectors. We define $T(M)$ as the quotient space of $\bigcup_\alpha U_\alpha \times R^n$ (disjoint union) by the equivalence relation

$$(h_\alpha^{-1}(x), y) \sim (h_\beta^{-1}(x), J(h_\beta^{-1}h_\alpha)y).$$

Then $T(M)$ is a vector bundle over M, where

$$p[(h_\alpha^{-1}(x),y)] = x, \text{ and } s(x) = [(h_\alpha^{-1}(x),0)],\ x \in V_\alpha;$$

and

$$k_\alpha(x,y) = [(h_\alpha^{-1}(x),y)],\ x \in V_\alpha,\ y \in R^n.$$

Note that this definition enables us to talk about tangent vectors to smooth curves in M. For a smooth surface in R^{n+k}, if we identify the tangent vector to $x(t)$ in M to the tangent vector to $x(t)$ in R^{n+k}, we see that $T(M)$ gets identified to the tangent bundle to the surface as previously defined.

Unfortunately, we are still not in a position to generalize the notion of tangent bundle to topological manifolds since our definition still involves derivatives. We need a representation of the tangent bundle of a smooth manifold which can be specified without explicit reference to derivatives. For this purpose we shall use the fact (see below) that for a Riemannian manifold the tangent bundle may be embedded in $M \times M$ as a neighborhood of the diagonal

$$\Delta(M) = \{(x_1, x_2) \in M \times M \mid x_1 = x_2\}.$$

Now any smooth manifold may be given a Riemannian metric (an inner product on $T_x(M)$ for each $x \in M$, which depends smoothly on M). This may be done by using the standard inner product (dot product) on R^n for each $U_\alpha \times R^n$ and then piecing these inner products together by a smooth partition of unity on M, subordinate to the cover $\{U_\alpha\}$. (This uses the fact, easily deduced, that a manifold is paracompact.) This defines the notion of length for tangent vectors, and by integration, a length for piecewise smooth curves. This enables one to write down a differential equation for curves of shortest length; i.e., geodesics, between nearby points. The solution is unique and depends smoothly on the endpoints. This in turn enables one to define a smooth map called the exponential, $\exp_x: T_x(M) \to M$, sending a tangent vector $v \in T_x(M)$ onto the geodesic of length $\| v \|$ issuing from x in the direction v (i.e., tangent vector of the geodesic at x is v). Because of the uniqueness of geodesics between nearby points, there is a neighborhood V of the zero vector in $T_x(M)$, such that $\exp_x | V$ is one to one onto a neighborhood W of $x \in M$; and one may show it has a smooth inverse.

Thus we may define an embedding ϕ of a neighborhood of the zero section of $T(M)$ onto a neighborhood of $\Delta(M) \subset M \times M$ by $\phi(x,v) = (x, \exp_x v)$. By shrinking each fibre $T_x(M) \cong R^n$ radially into itself, we can define a smooth fibrewise embedding r of $T(M)$ in any neighborhood of the zero section s. Then $\psi = \phi \circ r: T(M) \to M \times M$ embeds $T(M)$ as a neighborhood of the diagonal, and satisfies:

(a) $\mathrm{proj}_1 \circ \psi = p$, $\mathrm{proj}_1: M \times M \to M$

(b) $\psi \circ s(x) = (x,x) \in \Delta(M)$.

Conversely one has:

PROPOSITION 3.1: *Let $M \xrightarrow{s} E \xrightarrow{p} M$ be a smooth n-dimensional vector bundle and $\psi: E \to M \times M$ a smooth embedding onto a neighborhood of $\Delta(M)$ satisfying* (a) *and* (b) *above; then (s,p) is linearly equivalent to $T(M)$.*

Sketch of Proof. On the one hand, the vectors tangent to the fibres $p^{-1}(x)$ in E at the zero section form a bundle which may be identified to the given bundle (s,p). This may be seen by using the local product structure on E.

On the other hand, ψ maps $p^{-1}(x)$ smoothly onto a neighborhood of (x,x) in $x \times M \subset M \times M$. But the vectors in $M \times M$ tangent to $x \times M$ at (x,x) can obviously be identified to $T_x(M)$, and it is not difficult to see that the set of all such

"vertical vectors" in $T(M \times M)$ along $\Delta(M)$ is a vector bundle which can be identified to $T(M)$. Thus ψ defines a linear equivalence of (s, p) and $T(M)$.

4. The tangent bundle of a topological manifold [4]. Proposition 3.1 suggests the

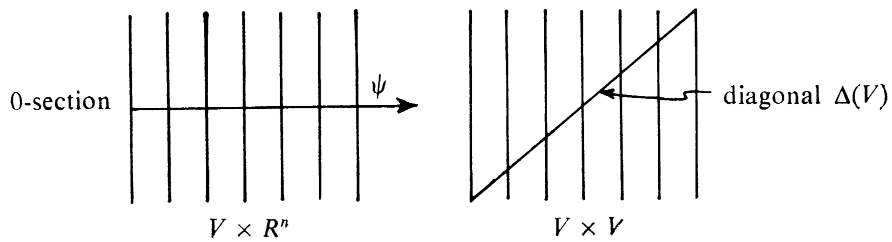
DEFINITION 4.1: Let M^n be a topological n-manifold. An R^n-bundle $M \xrightarrow{s} E \xrightarrow{p} M$ is called a **tangent bundle** of M if there exists an embedding $\psi: E \to M \times M$ sending E onto a neighborhood of the diagonal and satisfying (a) and (b) above.

PROPOSITION 4.2: *Let M^n be a topological manifold, then M has a tangent R^n-bundle, unique up to equivalence.*

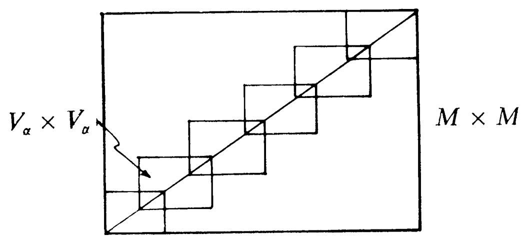
Sketch of Proof. We first show that for each $x \in M$, there is a neighborhood V such that V has a trivial tangent bundle. Let $x_0 \in M$, and let $h: U \to V$ be a homeomorphism of an open set in R^n with an open neighborhood V of x_0 in M. Since $h^{-1}(x_0)$ has an open ε-neighborhood in U, and such an open ε-ball is homeomorphic to R^n, we may assume h is a homeomorphism of R^n onto the open neighborhood V of x_0, with $h(0) = x_0$. Then the homeomorphism

$$\psi: V \times R^n \xrightarrow{h^{-1} \times id} R^n \times R^n \xrightarrow{\phi} R^n \times R^n \xrightarrow{h \times h} V \times V \subset M \times M,$$

$$\phi(z, y) = (z, z + y), \text{ satisfies (a) and (b).}$$



Unfortunately, these local products do not fit together by homeomorphism, rather we have the following picture:



However, one can show that such a neighborhood of the diagonal does contain

an R^n-bundle, unique up to equivalence, by inductively evening out the overlaps, using [2]:

KISTER'S THEOREM. *Let* $\varepsilon_0(R^n, R^n)$ *be the space of embeddings of* R^n *into* R^n *sending zero into zero, with the compact open topology. Let* $H_0(R^n)$ *be the subspace of* $\varepsilon_0(R^n, R^n)$ *consisting of homeomorphisms. Then* $H_0(R^n)$ *is a (weak) deformation retract of* $\varepsilon_0(R^n, R^n)$.

Using the tangent bundle of a topological manifold we can put the results on smoothing topological manifolds in a particularly nice form. (Results on triangulating topological manifolds, or smoothing piecewise linear manifolds can be put in a similar form once the tangent bundle of a piecewise linear manifold has been defined.)

Note that if a topological manifold M admits a smooth structure, the tangent vector bundle of the smooth structure embedded as a neighborhood of the diagonal (w.r.t. some Riemannian structure) represents a tangent R^n-bundle of M by ignoring the linear structure. In particular, this says that if M admits a smooth structure its tangent R^n-bundle admits a linear structure.

One has the partial converse [3]:

THEOREM 4.3: *If* M^n *is non-compact, and if the tangent* R^n*-bundle of* M *admits a linear structure; then* M *admits a smooth structure.*

If $n \geqq 5$, then we have stronger results. First we need:

DEFINITION 4.4. Let M_1 and M_2 be two smooth manifolds with the same underlying topological manifold M. Then M_1 and M_2 are called **isotopic smoothings** of M if $id_M\colon M_1 \to M_2$ is homotopic through homeomorphisms to a diffeomorphism.

DEFINITION 4.5: Let (s_1, p_1) and (s_2, p_2) be two vector bundles with the same underlying R^n-bundle (s, p). Then (s_1, p_1) and (s_2, p_2) are called **isotopic linearizations** of (s, p) if $id_E\colon E_1 \to E_2$ is homotopic through bundle equivalences to a linear bundle equivalence.

THEOREM 4.6: *If* $n \geqq 5$, *and* M^n *is any topological manifold, the isotopy classes of smoothings of* M *are in* 1-1 *correspondence with the isotopy classes of linearizations of the tangent bundle of* M.

These results can in turn be expressed in terms of algebraic topology: Let $\mathrm{Top}_n = H_0(R^n)$ and O_n = orthogonal transformations of R^n. Then $O_n \subset \mathrm{Top}_n$ and we can form the homogeneous space Top_n/O_n.

It follows that the obstructions to deforming one smoothing of M to another lie in the cohomology of M with coefficients in the homotopy groups of Top_n/O. Also, the obstructions to smoothing M can be similarly expressed.

References

1. R. L. Bishop and R. J. Crittenden, Geometry of Manifolds, Academic Press, New York, 1964.
2. J. Kister, Microbundles are fibre bundles, Ann. Math., (2) 80 (1964) 190–199.
3. R. Lashof, The immersion approach to triangulation and smoothing, Proc. Symp. in Pure Math., Amer. Math. Soc., XXII (1971) 131–164.
4. J. Milnor, Microbundles, Proc. Int. Cong. Math. (Stockholm 1962), Inst. Mittag-Leffler,. Djursholm 1963.
5. R. Schultz, Some recent results on topological manifolds, this MONTHLY, 78 (1971) 941–951.
6. N. Steenrod, The Topology of Fibre Bundles, Princeton University Press, Princeton, N. J., 1951.

MORE ON THE SUPERPARTICULAR RATIOS IN MUSIC

G. D. HALSEY AND EDWIN HEWITT, University of Washington

There are ratios that are assigned without hesitation to the musical intervals that are the basis of traditional Western music. That is, these ratios denominate the relative acoustic frequency, or inversely, the length of violin string required to produce first one note and then the other of the interval. A recent article in this MONTHLY [6] by A. L. Leigh Silver presents an interesting discussion of this fact, and lists the following *superparticular* ratios along with their proper musical designations:

$2/1$ octave	$9/8$ major whole tone
$3/2$ perfect fifth	$10/9$ minor whole tone
$4/3$ perfect fourth	$16/15$ diatonic semitone
$5/4$ major third	$25/24$ chromatic semitone
$6/5$ minor third	$81/80$ common comma [or comma of Didymus].

In essence, these designations appear to have been known since the times of Zarlino and Descartes [2, p. 775]. With inversions, they account for all the common intervals except the tritone. The unstable character of the tritone sets it apart, as discussed, for example, by Hindemith [3, p. 81]. It can be expressed as a ratio by compounding suitable superparticular ratios. Whether it is assigned the ratio $64/45$ or $45/32$, depending on musical context, or indeed some other ratio, it is not superparticular, which is in keeping with its unique rôle in music.

Silver implies that the above ratios, limited to contain prime factors of 2, 3 and 5, are a finite sequence. It has been long known that the sequence actually terminates with $81/80$: this was proved in 1897 by C. Størmer [7]. Størmer also proved a more

general theorem [**8**], as follows. (We are indebted to Professor Ivan Niven for this reference.) *Let* A, B, $M_1, \cdots, M_m$, N_1, $N_2, \cdots, N_n$ *be given positive integers. Then the equations*

$$AM_1^{x_1}M_2^{x_2}\cdots M_m^{x_m} - BN_1^{y_1}N_2^{y_2}\cdots N_n^{y_n} = \pm 1, \ \pm\ 2$$

admit only a finite number of solutions, all of which can be computed from the smallest positive solutions u_k *of Pell's equation*

$$t_1^2 - D_1u_1^2 = 1, \cdots, t_r^2 - D_ru_r^2 = 1$$

for certain D_k*'s that can be written down in terms of* A, B, M_i, *and* N_j.

D. H. Lehmer [**4**] has recently given a new proof of Størmer's theorem for prime M_j's and $A = B = 1$ (excluding ± 2 on the right side) and has published complete tables for the primes 2, 3, 5, ⋯, 41. (Professor Donald R. Snow has kindly given us this reference.)

It may be of some interest to give a short derivation of Størmer's theorem for our case. The pairs of integers $(x, x+1)$ for which x and $x+1$ are divisible only by 2, 3, or 5 are (1,2), (2,3), (3,4), (4,5), (5,6), (8,9), (15,16), (24,25), (80,81). That is to say, all possible superparticular ratios derived from the first three primes were long ago identified by musical theory.

We establish the result by checking all possible cases. We first note that if

$$(*) \qquad 2^a3^b5^c - 2^{a'}3^{b'}5^{c'} = \pm 1,$$

(all exponents nonnegative integers), then $aa' = bb' = cc' = 0$, since the left side has absolute value at least $2^{|a-a'|}$ if $a \neq a'$, for example. A moment's thought shows that the only possible solutions of the equation (*) are the following, where a, b, c denote positive integers:

(0) $1 = 2^1 - 1$;

(1) $2^a = 3^b \pm 1$; (2) $2^a = 5^b \pm 1$;

(3) $3^a = 5^b \pm 1$; (4) $2^a3^b = 5^c \pm 1$;

(5) $2^a5^b = 3^c \pm 1$; (6) $3^a5^b = 2^c \pm 1$.

For the two equations (1), we know the solutions $(a, b) = (1,1), (1,0), (2,1), (3,2)$. We shall show that there are no others. Assuming that there are other solutions, we may suppose that $a > 3$ and that a is the least value that yields a solution of the equations in (1). Plainly we have $b > 2$, and so $2^a \equiv \pm 1 \pmod 9$. Since $2^a \equiv 1 \pmod 9$ if and only if $a \equiv 0 \pmod 6$ and $2^a \equiv -1 \pmod 9$ if and only if $a \equiv 3 \pmod 6$, it follows that $a \equiv 0 \pmod 3$: $a = 3a'$. Thus we have

$$2^{3a'} \pm 1 = (2^{a'} \pm 1)x = 3^b$$

for some positive integer x, and so unique factorization shows that $2^{a'} \pm 1 = 3^{b'}$.

The minimum condition on a and the restriction $a > 3$ show that $a' = 2$ or 3, i.e., $a = 6$ or 9. Since $2^6 \pm 1 = 63, 65$ and $2^9 \pm 1 = 511, 513$, we see that (1) admits no solutions besides those listed above.

For the two equations (2), we know one solution, namely $(a, b) = (2,1)$. If there are others, suppose that we have the least exponent $b > 1$. Thus we have $5^b \equiv \pm 1 \pmod 8$, and since $5^{2k+1} \equiv 5 \pmod 8$, $5^{2k} \equiv 1 \pmod 8$, we see that $2^a = 5^b + 1$ has no solution. For $2^a = 5^b - 1$, we get $2^a = 5^{2b'} - 1$, $(5^{b'} + 1)\ (5^{b'} - 1) = 2^a$. Now argue as in the discussion of equations (1).

The equations (3) trivially have no solutions since one side is even and the other is odd.

In the case of equations (4) consider the equation $2^a 3^b = 5^c - 1$. We know the solution $(a, b, c) = (3,\ 1,\ 2)$. Plainly we must have $c > 1$, and so

$$2^a 3^b = 2^2 \sum_{j=0}^{c-1} 5^j,$$

which implies that $a \geqq 2$. Since $\sum_{j=0}^{c-1} 5^j \equiv 0 \pmod 3$, c has to be even, $c = 2c'$, and we have

$$2^a 3^b = (5^{c'} - 1)(5^{c'} + 1).$$

The number $5^{c'} + 1$ is congruent to 2 modulo 4. Unique factorization and the last equality yield

$$5^{c'} + 1 = 2 \cdot 3^{b'},\ 5^{c'} - 1 = 2^{a-1} \cdot 3^{b-b'}$$

for some integer b' such that $1 \leqq b' \leqq b$. Subtracting, we find

$$1 = 3^b - 2^{a-2} \cdot 3^{b-b'}.$$

Plainly we must have $a > 2$, and also either $b' = 0$ or $b = b'$. Since $b' \geqq 1$, we have

$$5^{c'} - 1 = 2^{a-1},$$

which by the above solution of (2) implies that $a - 1 = 2$, $c' = 1$. Thus $2^3 3^1 = 5^2 - 1$ is the only solution of $(4-)$.

Next consider the equation $2^a 3^b = 5^c + 1$, for which we know the solution $2^1 3^1 = 5^1 + 1$. Assuming that there is a solution with $c > 1$, we may suppose that we have the solution with the least value of $c > 1$. Since the right side is congruent to 2 modulo 4, we must have $a = 1$. Since $2 \cdot 3^b \equiv 1 \pmod 5$ if and only if $b \equiv 1 \pmod 4$, we have $b = 4b' + 1$ with $b' \geqq 0$. Since $5^c \equiv 1 \pmod 3$ if and only if c is odd, we have $c = 2c' + 1$, with $c' \geqq 0$, and so our equation is

$$\begin{aligned} 2 \cdot 3^{4b'+1} &= 5^{2c'+1} + 1 \\ &= 6 \sum_{j=0}^{2c'} (-1)^j 5^j, \end{aligned}$$

i.e.,

$$3^{4b'} = \sum_{j=0}^{2c'} (-1)^j 5^j.$$

If $b' = 0$, we have $c' = 0$ and we are at our known solution $a = b = c = 1$. If $b' > 0$, we argue as follows. Since $-5 \equiv 1 \pmod 3$, we have

$$\sum_{j=0}^{2c'} (-1)^j 5^j \equiv 2c' + 1 \pmod 3,$$

and so $2c' + 1 \equiv 0 \pmod 3$. That is, c has the form $3(2d + 1)$, and our original equation has the form

$$\begin{aligned} 2 \cdot 3^{4b'+1} &= 5^{3(2d+1)} + 1 \\ &= (5^{2d+1} + 1)(5^{2(2d+1)} - 5^{2d+1} + 1). \end{aligned}$$

Applying unique factorization, we see that there is a b'' such that

$$2 \cdot 3^{4b''+1} = 5^{2d+1} + 1.$$

Since $c = 3(2d + 1)$ is the least value of $c > 1$ yielding a solution of (4+), we see that $2d + 1 = 1$, $c = 3$. Since $5^3 + 1 = 2 \cdot 3^2 \cdot 7$, we have proved that (4 +) has only one solution, $2^1 \cdot 3^1 = 5^1 + 1$.

For the equation (5), we have only the solutions $(a, b, c) = (4, 1, 4)$ and $(1, 1, 2)$.

The equation (6−): $3^a 5^b = 2^c - 1$ has only the solution $(1, 1, 4)$ and the equation (6 +): $3^a 5^b = 2^c + 1$ has no solutions at all. The proofs are like those gone through above and are omitted.

Although ratios that involve the number 7 are foreign to the true musical intervals, in at least one instance, Hindemith [loc.cit. p. 82] uses two such ratios in a tentative analysis of the dominant seventh chord. There he ascribes the ratios 7/5 or 10/7 to the tritone. Although these ratios are not superparticular, the interval that characterizes their difference (50/49) is superparticular. Therefore, it is of some mild interest for musical theory to list the solutions of Størmer's equation for the primes {2,3,5,7} and $\pm$ 1. A computation yields: (6,7), (7,8), (14,15), (20,21), (27,28), (35,36) (48,49), (49,50), (63,64), (125,126), (224,225), (2400,2401), (4374,4375). Størmer [7] has shown that these are the only adjacent pairs for the primes 2,3,5,7. Lehmer [4] has a complete table for the primes 2,3,5,$\cdots$,41.

There is a generalization of part of Størmer's theorem, which follows readily from a theorem of A. Baker [1]. (We are indebted to Professor James Jordan for the reference to Baker's article.) Given any finite set P of primes and any fixed positive integer a, there are only a finite number of pairs (x, y) of positive integers such that $|x - y| \leqq a$ and x and y admit as prime factors only numbers from P.

Finally we note the interesting paper of Pólya [5], where analogues of part of Størmer's theorem are taken up.

References

1. A. Baker, Linear forms in the logarithms of algebraic numbers (IV). Mathematika, 15 (1968) No. 30, 204–216.

2. Grove's Dictionary of Music and Musicians, 3rd edition, Vol. V. H. C. Colles, editor. Macmillan, New York, 1936.

3. Paul Hindemith, The Crafts of Musical Composition, Book I. Associated Music Publishers, New York, 1945.

4. D. H. Lehmer, On a problem of Størmer. Illinois J. Math., 8(1964) 57–79.

5. George Pólya, Zur arithmetischen Untersuchung der Polynome, Math. Z., 1(1918) 143–148.

6. A. L. Leigh Silver, Musimatics or the nun's fiddle, this MONTHLY, 78(1971) 351–357.

7. Carl Størmer, Quelques théorèmes sur l'équation de Pell $x^2 - Dy^2 = \pm 1$ et leurs applications, Skrifter Videnskabs-selskabet (Christiania) I, Mat.-Naturv. Kl., no.2 (1897) 48p.

8. ———, Sur une équation indéterminée, C. R. Acad. Sci. Paris, 127 (1898) 752–754.

CORRECTION TO "RECONSTRUCTING AN EVOLUTIONARY TREE"

(This MONTHLY, 79(1972), 596–603)

DAVID SANKOFF

The figure on p. 597 should be labelled FIG. 2 and should appear on p. 600; the figure which appears on p. 600 should be labelled FIG. 1a and FIG. 1b and should appear on p. 597.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

COMPLEMENTS AND COMMENTS

ROBERT GILMER

We are grateful to readers who are willing to share with us their comments on articles appearing in the Notes Section. Such comments enhance the value of the Monthly. The information we have received during the past year includes the following.

Calculus. J. D. Riley notes that the necessary hypothesis $f(0) = 0$ has been omitted in the article by F. Cunningham and N. Grossman (September, 1971, pp. 781–3) concerning Young's inequality.

In an article in the June–July 1972 MONTHLY, pp. 634–5, G. J. Porter calls attention to the Cauchy Condensation Test as an alternative to the integral test for convergence of infinite series. D. Drasin and Ralph Garfield point out that the material in Porter's article is contained in pages 46–48 of the popular text *Principles of Mathematical Analysis*, by W. Rudin.

Concerning the same article, Ray Glenn notes the following two minor errors. In the proof of the Cauchy Condensation Test, the last inequality should be the equality $\cdots = a_1 + 2\sum_{j=1}^{\infty} 2^{j-1} a_{2j}$, and the final expression in Example 3 should be

$$\frac{1}{\log 2} \sum \frac{1}{j[\log j + \log(\log 2)]^{\alpha}}.$$

General. M. R. Sridharan expresses objections, based on mathematical logic and language, to H. C. Kennedy's paper concerning Boyer's law (January, 1972, pp. 66–7). Sridharan does not consider Kennedy's statement of Boyer's law to be, in itself, a mathematical law or theorem, and hence he takes exception to the assertion that Boyer's law is a rare instance of a law whose statement confirms its own validity. Even with a restatement of Boyer's law, such as "discoveries are not usually attributed to their original discoverers," Sridharan still objects to the inclusion of the word "usually."

Geometry and Topology. Murray Klamkin has indicated to us a shorter, more geometric proof of the theorem in J. C. C. Nitsche's paper *The smallest sphere containing a rectifiable curve* (October, 1971, pp. 881–2).

D. E. Sanderson writes that one of his former students, A. Irudayanathan, proved the following result in his unpublished 1967 Iowa State University Ph.D. Thesis. If Ω is the collection of all open covers of a space Y and for $\alpha \in \Omega$, $y \in Y$, $\alpha^*(y)$ and $\alpha^{**}(y)$ are defined by $\alpha^*(y) = \cup\{U \in \alpha \mid y \in U\}$ and $\alpha^{**}(y) = \cup\{\alpha^*(x) \mid x \in \alpha^*(y)\}$, then Y is regular if and only if $\{\alpha^{**}(y) \mid y \in Y, \alpha \in \Omega\}$ is a basis for the topology of Y. A similar observation is contained in the June–July 1972 article by James Chew (pp. 630–2). Sanderson also points out that the open cover U in Chew's condition III is superfluous—only an open set U containing the point a is used.

Set theory and logic. P. G. J. Vredenduin (1969, 59) has given an alternative to Russell's Paradox in classical set theory. S. K. Bose notes the similarities in the two examples, and expresses the opinion that, in fact, Russell's Paradox may have pedagogical advantages, since Vredenduin requires the fact that $\# X < \# \mathcal{P}(X)$ for each set X.

Algebra. Two readers, A. S. Fraenkel and G. Haggard, have pointed out that the theorem proved in Emile Roth's article (November, 1971, pp. 990–2) concerning permutations arranged around a circle is contained in a 1946 paper of I. J. Good (J. London Math. Soc., Vol. 21, pp. 167–9). Haggard also calls attention to the fact that Chapter 9 of Sherman Stein's book *Mathematics—the Man-made Universe* is devoted to an exposition of this problem, its solution, and relevant bibliography.

The article by R. L. Roth (1971, pp. 392–3) on extensions of the rationals by square roots continues to attract the interest of our readers (see p. 1105 of the December,1971 MONTHLY).In a paper entitled *On the linear independence of algebraic numbers*, Pacific J. Math., 3 (1953) 625–630, L. J. Mordell (now deceased) proved the following result. Let K be an algebraic number field, let $a_1, \cdots, a_s$ be elements of K, and let $n_1, \cdots, n_s$ be positive integers. For $1 \leqq i \leqq s$, let t_i be a complex root of the polynomial $X^{n_i} - a_i$. If $P(X_1, \cdots, X_s) \in K[X_1, \cdots, X_s]$ is a nonzero polynomial of degree less than n_i in X_i for each i, and if there exists no relation of the form $t_1^{e_1} t_2^{e_2} \cdots t_s^{e_s} = a$, where $a \in K$, unless $e_i \equiv 0 \pmod{n_i}$ for each i, then $P(t_1, \cdots, t_s) \neq 0$. Roth's result is, of course, a special case of the preceding theorem of Mordell. In fact, Roth's theorem follows from a less general form of Mordell's theorem due to A. Besikovitch (J. London Math. Soc., 15 (1940) 3–6).

Robert MacKenzie and John Schuneman (October, 1971, pp. 882–3) give an example of a finite algebraic number field F and a quadratic extension K of F such that K/F does not have a relative integral basis. William C. Waterhouse observes that such examples abound, in view of the following theorem of H. B. Mann (Proc. Amer. Math. Soc., 9 (1958) 167–172). If D is a Dedekind domain, not a principal ideal domain, with quotient field L, then there exists a quadratic extension field L' of L such that D', the integral closure of D in L', is not a free D-module (or, in other language, L'/L does not have an integral basis).

According to Donald J. McCarthy, the characterization of supersolvable groups published by W. E. Deskins (1968, pp. 180–2) has a previous history, as described on page 590 of McCarthy's survey article in Transactions of the New York Academy of Sciences, Series II, 33(1971), 586–594.

Number theory. M. J. DeLeon writes about a slight strengthening of the results announced by D. A. Butter (December, 1971, p. 1109). By observing not only that $x^p - x \equiv 0 \pmod 2$, but $x^p - x \equiv 0 \pmod 6$, the conclusion of the theorem (under the assumption that $p > 3$) can read "$x_1 + \cdots + x_n \geqq z + 6p$ and $p < (n-1)z/6$." Also, the conclusion of the corollary can read "$p < \min\{x, y\}/6$".

Volume 74 of the MONTHLY contained two articles concerning Farey series that referred to a paper by Jean Blake (1966, pp. 50–2). The articles in question were by Alan Zame (October, 1967, p. 977) and by Irving Katz (December, 1967, p. 1233); Kim Ki-Hang Butler has noted the close relationship between the results of Zame and of Katz.

If p is prime, if n is a positive integer, and if k is a non-negative integer such that p^k divides n but p^{k+1} does not divide n, then G. J. Simmons (1970, pp. 510–1) denotes this relationship by the symbol $p^k \| n$. Simmons proved that if r is a positive integer and if $p_1, \cdots, p_k$ are arbitrary primes, then there are infinitely many positive integers n such that $p_j^0 \| \binom{n}{r}$ for each j between 1 and k. F. T. Howard has extended Simmons'

theorem to the case where not only are r and $p_1, p_2, \cdots, p_k$ specified in advance, but also non-negative integers $i_1, \cdots, i_k$ are specified; the conclusion is that there are infinitely many positive integers n such that $p_j^{i_j} \| \binom{n}{r}$ for each j between 1 and k. Note that the 1970 *Complements and Comments* article (p. 1078) also contained some remarks on Simmons' paper.

Arthur Marshall points out that the article of C. Vanden Eynden (June–July, 1972, p. 625) implicitly contains a "formula" for the mth prime, for $m \geqq 2$: If $p_1, \cdots, p_{m-1}$ are the first $m-1$ primes, and if $Q = p_1 \cdots p_{m-1}$, then the mth prime is d_2, where $\{d_1 < d_2 < d_3 < \cdots < d_{\phi(Q)}\}$ is the set of positive integers less than Q that are relatively prime to Q; alternately, the mth prime is $Q - d_{\phi(Q)-1}$.

Analysis. In the October, 1971 MONTHLY, W. R. Bauer and R. H. Benner present a proof, independent of category theory, of the result that a Hamel basis for an infinite dimensional Banach space has cardinality greater than $\aleph_0$. Bauer-Benner state that textbooks either omit the result or defer it until after some category theory has been developed, but William R. Transue notes that this is not quite the case—in Problem 2, page 109 of J. Dieudonné's *Foundations of Modern Analysis* (where the category theorem is not proved at all) a proof of the theorem cited is sketched along lines similar to those used by Bauer and Brenner.

James S. Byrnes points out that in regard to his article in the May, 1972 MONTHLY, pp. 510–2, a trivial example of a complete sequence that is not a basis can be obtained by adjoining any L^2 function to a given basis; A. Wilansky has communicated to us the same observation. The example in Byrnes' paper is nontrivial in the sense that there are functions f in L^2 for which no sequence $\{a_n\}$ such as is described in the definition of a basis exists; in particular, $f(x) \equiv 1$ is such a function. Wilansky also points out that by using the notion of a biorthogonal system, Byrnes could have greatly simplified the proof that the sequence in his example is not a basis.

R. P. Boas has sent us comments concerning two of his own articles that appeared in Volume 78 of the MONTHLY. In his article on signs of derivatives and analytic behavior (pp. 1085–1093), Boas failed to cite B. McMillan, Ann. of Math., 60 (1954) 467–501, a paper containing the widest generalization to date. With regard to his paper with J. W. Wrench (pp. 864–870), Boas states that the book *Matter, Earth, and Sky*, by G. Gamow, Prentice Hall, 1958 contains on pages 15, 16 some material on S_n disguised as the problem of how far the top book of a stack of n identical books can be made to project beyond the edge of the table on which the stack rests (the distance is $\frac{1}{2} S_n$). Gamow's intuition failed him, however, since he says "Because of the rapidly decreasing contribution of each new book, however, we will need the entire Library of Congress to make overhang equal to three or four books lengths!" His problem is to make S_n exceed 6 or 8, and by the tables in the paper of Boas and Wrench, this already occurs with 227 or 1674 books.

DIVERGENCE CRITERIA FOR POSITIVE SERIES

D. BORWEIN, University of Western Ontario, and A. MEIR, University of Alberta

Suppose throughout that f is a mapping of the set of positive integers into itself, and that $\{\lambda_n\}$ is a sequence of real non-negative numbers.

Using a combinatorial argument, K. A. Post [1] recently established the following result:

Let

(1) $$f(n+1) - f(n) \geqq n + 1, \quad n = 1, 2, \cdots,$$

and suppose that the sequence $\{a_n\}$ *satisfies*

(2) $$0 < a_n \leqq a_{n+1} + a_{f(n)}, \quad n = 1, 2, \cdots.$$

Then $\Sigma\, a_n = \infty$.

Post notes Erdös's observation that if $f(n) \leqq cn^2$, $0 < c < \frac{1}{2}$, then (2) does not, in general, imply the divergence of $\Sigma\, a_n$.

Our first theorem extends the scope of the above divergence criterion by showing that (1) and (2) can be replaced by more general inequalities.

THEOREM 1. *Let*

(3) $$\lambda_n \leqq 1, \qquad n = 1, 2, \cdots,$$

(4) $$f(n+1) - f(n) \geqq n\lambda_n + 1, \quad n = 1, 2, \cdots,$$

and suppose that the sequence $\{a_n\}$ *satisfies*

(5) $$0 < a_n \leqq a_{n+1} + \lambda_n a_{f(n)}, \quad n = 1, 2, \cdots.$$

Then $\Sigma\, a_n = \infty$.

Our second theorem shows that for a decreasing sequence $\{a_n\}$, condition (3) of Theorem 1 is redundant when a slightly modified version of condition (4) holds.

THEOREM 2. *Let*

(6) $$f(n+1) - f(n) \geqq n\lambda_{n+1} + 1, \quad n = 1, 2, \cdots,$$

and suppose that $\{a_n\}$ *is a decreasing sequence satisfying* (5). *Then* $\Sigma\, a_n = \infty$.

REMARKS. (i) Post's combinatorial argument cannot be used in the proof of Theorem 1 because, in general, we shall have $\lambda_n < 1$ for some values of n.

(ii) Condition (6) cannot be replaced by (4) in Theorem 2. Indeed, if $\{a_n\}$ is *any* given sequence of positive numbers, we can define f and $\{\lambda_n\}$ by induction so that (4) and (5) hold: Let $f(1) = 1$ and suppose $\lambda_1, \lambda_2, \cdots, \lambda_{m-1}$ and $f(1), f(2), \cdots, f(m)$ are known. First define λ_m so that (5) holds for $n = m$, and then define $f(m+1)$ so that (4) is satisfied for $n = m$.

(iii) Conditions (4) and (5) alone do not, in general, imply the divergence of $\Sigma\, a_n$, even when λ_n is constant. This we can demonstrate by means of the following example: Let $\lambda_n \equiv \lambda > 1$ and let f satisfy (4). Let $f^0(m) = m$ and $f^\nu(m) = f(f^{\nu-1}(m))$. For $n = 1, 2, \cdots$, let $k = k_n$ be the largest non-negative integer for which $f^k(r) = n$. Having thus determined k, a simple argument shows that $r = r_n$ is also uniquely determined. We define $\{a_n\}$ by $a_n = \lambda^{-k}2^{-r}$. It is easily seen that (5) holds. But $\Sigma\, a_n \leqq \Sigma_{k,r}\lambda^{-k}2^{-r} < \infty$.

(iv) In neither Theorem 1 nor Theorem 2 can the coefficient λ_n in (5) be replaced, in general, by any larger number even for a decreasing sequence $\{a_n\}$. For, let $\lambda_n \equiv \lambda > 0$, let $a_n = 1/n \log n (\log\log n)^2$ for $n \geqq 2$, and let $f(n)$ be defined by $f(1) = 1$ and $f(n+1) - f(n) = 2 + [\lambda n]$ for $n \geqq 1$. Then, for arbitrary $\varepsilon > 0$, we have $a_n \leqq a_{n+1} + (\lambda + \varepsilon)a_{f(n)}$ if n is large enough. But $\Sigma\, a_n < \infty$.

Proof of Theorem 1. If $\lambda_n \equiv 0$ then $a_n \geqq a_1 > 0$ for all n, and the required conclusion is trivial. Assume therefore that $\lambda_{N-1} > 0$. Then by (4), $f(N) > N$.

Now, from (5) we have by iteration that, for $n \geqq 1$, $f(n) < m < f(n+1)$,

$$a_m \geqq \max\{a_{f(n)} - \sum_{r=f(n)}^{f(n+1)-1} \lambda_r a_{f(r)},\ 0\} = b_n$$

say, and hence, by (4), we have that

$$(7)\qquad \sum_{m=f(n)}^{f(n+1)-1} a_m \geqq a_{f(n)} + \{f(n+1) - f(n) - 1\}b_n \geqq a_{f(n)} + n\lambda_n b_n$$

$$\geqq a_{f(n)} + \sum_{k=N}^{n} \lambda_n b_n.$$

Suppose that $\Sigma\, a_m < \infty$. Summing on both sides of (7) for $N \leqq n < \infty$, we have, after interchanging the order of summations on the right, that

$$(8)\qquad \sum_{m=f(N)}^{\infty} a_m \geqq \sum_{n=N}^{\infty} a_{f(n)} + \sum_{k=N}^{\infty}\sum_{n=k}^{\infty} \lambda_n b_n.$$

From the definition of b_n and since $\lambda_n \leqq 1$, it follows that

$$(9)\qquad \sum_{n=k}^{\infty} \lambda_n b_n \geqq \sum_{n=k}^{\infty} \lambda_n a_{f(n)} - \sum_{r=f(k)}^{\infty} \lambda_r a_{f(r)} = \sum_{n=k}^{f(k)-1} \lambda_n a_{f(n)}.$$

Further, by (5),

$$\sum_{n=k}^{f(k)-1} \lambda_n a_{f(n)} \geqq a_k - a_{f(k)},$$

whence, from (8) and (9), $\Sigma_{m=f(N)}^{\infty}\, a_m \geqq \Sigma_{k=N}^{\infty}\, a_k$. The last inequality is impossible, since $f(N) > N$ and $a_N > 0$. Therefore $\Sigma\, a_n = \infty$.

Proof of Theorem 2. As in the proof of Theorem 1, we may assume that $\lambda_{N-1} > 0$,

so that $f(N) > N$. From (5) we get for $N \leqq n \leqq M$,

$$a_{f(n)} \leqq a_{f(M)+1} + \sum_{k=f(n)}^{f(M)} \lambda_k a_{f(k)}. \tag{10}$$

Multiplying both sides of (10) by $f(n)-f(n-1)$ and summing for $N \leqq n \leqq M$ we obtain

$$\begin{aligned} \sum_{n=N}^{M} \{f(n)-f(n-1)\}a_{f(n)} &\leqq f(M)\ a_{f(M)+1} + \sum_{n=N}^{M} \{f(n)-f(n-1)\} \sum_{k=f(n)}^{f(M)} \lambda_k a_{f(k)} \\ &\leqq f(M)a_{f(M)+1} + \sum_{k=f(N)}^{f(M)} \lambda_k a_{f(k)} \sum_{\substack{n \geqq N \\ f(n) \leqq k}} \{f(n)-f(n-1)\} \\ &\leqq f(M)a_{f(M)+1} + \sum_{k=f(N)}^{f(M)} \lambda_k a_{f(k)}\{k-f(N-1)\}. \end{aligned}$$

Now by (6), $\lambda_k\{k-f(N-1)\} \leqq \lambda_k(k-1) < f(k)-f(k-1)$, whence

$$\sum_{n=N}^{M} \{f(n)-f(n-1)\}a_{f(n)} < f(M)a_{f(M)+1} + \sum_{k=f(N)}^{f(M)} \{f(k)-f(k-1)\}a_{f(k)}. \tag{11}$$

Suppose now that $\Sigma\, a_n < \infty$. Then, since $\{a_n\}$ is a decreasing sequence,

$$na_n \to 0 \quad \text{as} \quad n \to \infty, \tag{12}$$

$$\sum_{n=N}^{\infty} \{f(n)-f(n-1)\}a_{f(n)} \leqq \sum_{n=N}^{\infty} \sum_{\nu=f(n-1)}^{f(n)-1} a_\nu \leqq \sum_{n=1}^{\infty} a_n < \infty. \tag{13}$$

Letting $M \to \infty$ in (11), we get, on account of (12) and (13), that

$$\sum_{n=N}^{\infty} \{f(n)-f(n-1)\}a_{f(n)} \leqq \sum_{k=f(N)}^{\infty} \{f(k)-f(k-1)\}a_{f(k)}.$$

But this is impossible, since $f(N) > N$. Therefore $\Sigma\, a_n = \infty$.

The following questions may be of interest:

Given an increasing integer-valued function g, what properties must f have in order that $0 < a_n \leqq a_{g(n)} + a_{f(n)}$ be a divergence criterion?

For what pairs of mappings g,f of the set of integers into itself is it true that, for some integer x, all values $f(x)$, $g(x)$, $f(f(x))$, $f(g(x))$, $g(f(x))$, $g(g(x))$, $f(f(f(x)))$, $\cdots$ are different?

The second question arises naturally in connection with Post's combinatorial lemma.

Reference

1. K. A. Post, A combinatorial lemma involving a divergence criterion for series of positive terms, this MONTHLY, 77(1970) 1085–1087.

DIFFERENTIABILITY AT A CORNER FOR A SOLUTION OF LAPLACE'S EQUATION

N. M. WIGLEY, University of Windsor, Canada

In the theory of elliptic partial differential equations, it is frequently assumed that a solution of a boundary value problem has derivatives of a given order right up to the boundary. If the boundary has a corner, such an assumption is, of course, unjustified (though not infrequently made). Using reflection or other tricks, one can sometimes show the existence of certain derivatives; and, at least for operators of the form $Lu = \Delta u + au_x + bu_y + cu = f$ with Dirichlet and/or Neumann boundary conditions on the arcs forming the corner, it is known that the solution is at least Hölder continuous at the corner [**1**].

In this note we wish to show that the question of what boundary conditions near the corner should offer what differentiability properties at the corner is not obvious. Setting $x + iy = re^{i\theta}$, let $0 < \alpha < 2$ and let D be the domain given by

$$0 < r < 1, \ 0 < \theta < \pi\alpha.$$

Let $\Delta u = 0$ in D, $u = 0$ for $\theta = 0, \pi\alpha$, $u = \sin(n\theta/\alpha)$ for $r = 1$, where n is a positive integer. If one seeks a solution in a small enough class of functions [**1**] then the solution exists and is unique; it is, in fact, $u = r^{n/\alpha}\sin(n\theta/\alpha)$.

Differentiability at the origin is thus determined by the magnitude of n/α and whether or not n/α is an integer. Thus a change in the boundary values on the "far off" arc $r = 1$ affects the differentiability of the solution at the origin.

Reference

1. N. M. Wigley, Mixed boundary value problems in domains with corners, Math. Z., 115 (1970) 33–52.

ON THE EXISTENCE OF PERIODIC AND UNBOUNDED SOLUTIONS OF LINEAR DIFFERENTIAL EQUATIONS WITH NON-NEGATIVE DAMPING

L. E. THOMAS, Saint Peter's College

Two important results concerning linear differential equations with constant coefficients are these:

(i) If a and b are positive constants, all solutions of $y'' + ay' + by = 0$ decay exponentially to zero; and

(ii) if a and b are positive constants and f is a continuous periodic function, then no solution of $y'' + ay' + by = f$ is unbounded.

It is customary to call a the *damping coefficient* and b the *restoring coefficient*.

In this note we consider equations of the same form in which the constants a and b are replaced by functions α and β. As is well known, conditions (i) and (ii) may fail to hold if there is no damping, that is if $\alpha \equiv 0$. As the main result of this note

shows, these conditions also fail to hold if α is nonnegative and has only isolated zeroes.

We provide a simple example of this fact, using an analysis which uses only the standard theory of linear differential equations and Floquet theory. The results might be used in the classroom to illustrate the power of the theory in obtaining information about the solutions of equations even when solutions may be impossible or inconvenient to find.

THEOREM I. *Let p and g be functions which satisfy the following conditions on some open interval I:*

(i) $p \in C^2(I)$ *(P is twice continuously differentiable on I).*

(ii) $g \in C^1(I)$ *(g is continuously differentiable on I).*

(iii) $p''(t) + g[p(t)] = 0$ *for all* $t \in I$.

Then p' satisfies the equation.

$$y'' + k(p')^2 y' + \left(\frac{d}{dp} g(p) + kg(p)p\right) y = 0 \text{ in } I, \tag{1}$$

where k is a constant and $' \equiv d/dt$.

Proof. The proof is by direct verification. If $y = p'$, then $y' = p'' = -g(p)$ and

$$y'' = p''' = -\frac{d}{dp} g(p)p'.$$

We have

$$-\frac{d}{dp} g(p)p' - k(p')^2 g(p) + \frac{d}{dp} g(p)p' + k(p')^2 g(p) = 0\,.$$

We have verified that $y = p'$ is a solution of equation (1).

REMARK: Theorem I also holds if k is a function continuous on I.

COROLLARY I. *Let p and g be functions which satisfy the following conditions on an open interval I:*

(i) $p \in C^2(I)$.

(ii) $g \in C^1(I)$.

(iii) $\frac{1}{2}\{p'(t)\}^2 + \int_0^{p(t)} g(\eta)d\eta \equiv c$, *a constant, on I.*

(iv) *p' has only isolated zeros, or p' is identically zero on I.*

Then p' satisfies equation (1) in I.

Proof. From (iii) we find by differentiation

$$\left.\begin{aligned} &p'(t)p''(t) + g[p(t)]p'(t) = 0, \\ &p'(t)\{p''(t) + g[p(t)]\} = 0, \end{aligned}\right\} \quad \text{for } t \in I.$$

If p' is identically zero, the conclusion of the corollary holds trivially. We wish to show that if p' is not identically zero then $p''(t) + g[p(t)] = 0$ for all $t \in I$. Thus, condition (iii) of Theorem I will be satisfied and the result of the corollary follows. For $t_0 \in I$, it suffices to show that even if $p'(t_0) = 0$ it is still true that $p''(t_0) + g[p(t_0)] = 0$. By condition (iv), there is a neighborhood N of t_0 such that $p'(t) \neq 0$ for all $t \in N$, $t \neq t_0$. Then $p''(t) + g[p(t)] = 0$ for $t \in N$, $t \neq t_0$. By the continuity of p'' and g, it must follow that $p''(t_0) + g[p(t_0)] = 0$. Since t_0 is arbitrary, this shows that the conditions of Theorem I are satisfied for all $t \in I$.

Since we have one solution of equation (1) we can find another by reduction of order. Another method is the following: Let u be a solution of equation (1) which is linearly independent of the known solution $y = p'$. The Wronskian $W(p', u)$ of p' and u may be found from Abel's formula to be

$$W(p',u)(t) = c\exp\left(-\int^t k[p'(s)]^2 ds\right),$$

where c is a non-zero constant which we may take as $c = 1$. This choice of c merely determines a specific solution u. An equation for u is, therefore,

$$p'u' - p''u = \exp\left(-\int^t k(p')^2\right). \tag{2}$$

This is a first order linear equation so its solution can always be represented as an integral.

An interesting special case occurs if we assume that p is a nonconstant periodic function. In this case the following result holds:

THEOREM II. *Let p and g satisfy the conditions of Theorem I, with the additional condition that p be a nonconstant periodic function. Let the constant k be positive. Equation* (1) *then has two* (*non-trivial*) *linearly independent solutions; one periodic, the other decaying exponentially to zero as $t \to \infty$.*

Proof. Suppose p has period T. One solution of equation (1) is $y = p'$, which is periodic. Since all of the coefficients appearing in equation (1) are periodic, Floquet theory applies. See Minorsky [**1**; pp. 127-133].

From Floquet theory we know that linearly independent solutions of equation (1) may be represented as

$$r_1(t)e^{h_1 t} \qquad \text{and} \qquad r_2(t)e^{h_2 t}, \tag{3}$$

where r_1 and r_2 are both periodic of period T and h_1 and h_2 are in general complex and satisfy

$$h_1 + h_2 \equiv \frac{1}{T}\int_0^T -k[p'(t)]^2 dt \qquad (\mathrm{mod}\ 2\pi i/T).$$

Since p' is a solution of the equation, we can take $h_1 = 0$ and $r_1 = p'$. Thus

(4) $$h_2 \equiv \frac{-1}{T} \int_0^T k[p'(t)]^2 dt \qquad (\bmod 2\pi i/T),$$

and so has nonnegative real part. Thus the second solution $r_2(t)e^{h_2 t}$ decays exponentially to zero.

As an example, consider the equation

(5) $$y'' + (k\sin^2 t)y' + (1 - k\cos t \sin t)y = 0.$$

This equation is seen to satisfy the conditions of Theorem I with $g(p) = p$. Thus, the equation $p'' + g(p) = 0$ becomes $p'' + p = 0$, which has $p(t) = \cos t$ as a solution. Thus $y(t) = p'(t) = -\sin t$ is a solution of (5).

In this case

$$h_2 \equiv \frac{-1}{2\pi} \int_0^{2\pi} k \sin^2 t \, dt \qquad (\bmod 2\pi i/T)$$

$$\equiv -\frac{k}{2} \qquad (\bmod 2\pi i/T).$$

We now turn our attention to the non-homogeneous equation

(6) $$y'' + k(p')y' + \left(\frac{d}{dp} g(p) + kg(p)p'\right) y = f, \qquad k > 0,$$

where f is a continuous function and p is periodic. If we let $p'(t)$ and $r(t)e^{-ht}$, $h > 0$, be two linearly independent solutions of the complementary homogeneous equation, a particular solution of the non-homogeneous equation may be written

(7) $$\int_0^t G(t,s)f(s)\, ds,$$

where

(8) $$G(t,s) = \frac{p'(s)r(t)e^{h(s-t)} - p'(t)r(s)}{U(s)},$$

$$U(s) = p'(s)\{r'(s) - hr(s)\} - r(s)p''(s).$$

We note that U is a periodic function which is never zero.

It is possible to choose a periodic forcing function f in such a way that the particular solution of equation (6) is unbounded, even though the homogeneous equation (1) has non-negative damping. Thus, "resonance" can occur in a damped system in this case. To demonstrate this we need only choose $f(t) = r(t)$. The solution may then be written

(9) $$r(t)e^{-ht} \int_0^t e^{hs} \frac{p'(s)r(s)}{U(s)} ds - p'(t) \int_0^t \frac{r^2(s)}{U(s)} ds.$$

Since p, r, and U are periodic, there exist positive numbers M and M' such that

$$\left|\frac{p'(s)r(s)}{U(s)}\right| < M \qquad \text{for all } s, \qquad \left|\frac{1}{U(s)}\right| > M' \quad \text{for all } s\,.$$

The inequality

$$\left| r(t)e^{-ht} \int_0^t \frac{p'(s)r(s)}{U(s)} e^{hs} ds \right| < \left| r(t) M e^{-ht} \int_0^t e^{hs}\, ds \right.$$
$$= \left| r(t) \right| M e^{-ht}(e^{ht} - 1)$$

shows that the first term of (9) is bounded. From the inequality

$$\left| p'(t)M' \right| \int_0^t r^2(s)\, ds < \left| p'(t) \int_0^t \frac{r^2(s)}{U(s)} ds \right|$$

we see that the second term is unbounded as $t \to \infty$. (This last inequality follows since $U(x)$ never changes sign.) We have thus shown that the particular solution of (6) is unbounded.

Unfortunately, it is usually very difficult to find an explicit form for r. Theoretically, r could be obtained from equation (2) with $u(t) = r(t)e^{-ht}$, but the quadratures necessary to obtain r are not in general tractable. In the case of equation (5), for example, the equation for r is

$$(-\sin t)r'(t) + (\cos t - h \sin t)r(t) = \exp\left(- \int^t k \sin^2 s\, ds \right).$$

Acknowledgement. The author expresses his appreciation to W. E. Boyce for his criticisms and suggestions concerning this paper.

Reference

1. N. Minorsky, Nonlinear Oscillations, Van Nostrand, Princeton, 1962.

A LEMMA ON PARTITIONS

DONALD KNUTSON, Columbia University

The object of this note is to prove the following lemma, important (indeed originally conceived) for the theory of projective algebraic varieties [**1**, p. 181].

LEMMA. *Let n be a positive integer. Let π be a partition of a multiple of $n!$ into parts $1, 2, \cdots, n$. Then π can be grouped into a sum of partitions of $n!$*

Equivalent statement: Given any equation $k \cdot n! = \sum_{i=1}^{n} b_i i$, with integers k, $b_i \geqq 0$, there exist integers a_i, $i = 1, 2, \cdots, n$ with $b_i \geqq a_i \geqq 0$ such that $n! = \sum_{i=1}^{n} a_i i$.

Proof. We prove the equivalent statement. For m a positive integer, let $m^{\#}$ denote the least common multiple of the integers $1,2,\cdots,m$. Clearly $m^{\#}$ divides $m!$, or, to put it crudely, if we add up enough copies of $m^{\#}$, we get $m!$ The idea of the proof is to extract from the given partition all subsums of size $n^{\#}$ obtainable from adding together just ones, or just twos, etc. The sum of the remaining terms is then in general smaller than $n!$

Specifically, consider the equation $k \cdot n! = \sum_{i=1}^{n} b_i i$, where we can of course assume that $k \geqq 2$. For each $i = 1,2,\cdots,n$, b_i can be written (uniquely) as $b_i = (n^{\#}/i)q_i + r_i$ for some integer $q_i \geqq 0$ and integer remainder r_i, $0 \leqq r_i < n^{\#}/i$. Hence

$$k \cdot n! = n^{\#} \sum_{i=1}^{n} q_i + \sum_{i=1}^{n} r_i i.$$

Observe now that the second term $\sum r_i i$ is less than $n \cdot n^{\#}$. But it is easy to prove that if $n \geqq 6$, then $n \cdot n^{\#} < n!$ Thus the first term must be greater than $n!$ so we can choose some q_i', $0 \leqq q_i' \leqq q_i$, $i = 1,\cdots,n$ so that $(n^{\#}) \sum_{i=1}^{n} q_i' = n!$ Let $a_i = (n^{\#}/i)q_i'$. Thus the lemma is proved for $n \geqq 6$.

The remaining cases are proved in an *ad hoc* manner.

Consider the case $n = 5$. We are given an equation $k \cdot 120 = b_1 + 2b_2 + 3b_3 + 4b_4 + 5b_5$. We can first assume that $b_1 < 120$, $2b_2 < 120$, $3b_3 < 120$, $4b_4 < 120$ and $5b_5 < 120$ or we would be trivially done. In finding the a_i, it can only make our task harder if we group the ones into pairs. Hence we can assume $b_1 \leqq 1$. Similarly, grouping the two into pairs, we can assume that $b_2 \leqq 1$.

We now try to find subsets of the numbers adding up to $5^{\#} = 60$. If we can find two such, we are done. If not, at least two of the numbers $3b_3$, $4b_4$, and $5b_5$ are less than 60. Hence, since they are multiples of 3, 4, and 5 respectively, the two must be less than 58. The third must be less than 120 (by the assumption above) so by the same reasoning, less than 118. Hence the sum $b_1 + 2b_2 + 3b_3 + 4b_4 + 5b_5 = k \cdot 120$ is less than $1 + 2 + 58 + 58 + 118 = 237 < 240$. Hence $k = 1$.

The case $n = 4$ is similar but easier. The cases $n = 3,2,1$ are trivial. ■

Questions:

(1) Is the lemma true with $n!$ replaced by $n^{\#}$?

(2) What can be said about the function $n^{\#}$? Is there a Stirling formula for it? What are the properties of the function $\sum_{n=0}^{\infty}(x^n/n^{\#})$?

Reference

1. D. Knutson, Algebraic Spaces, Lecture Notes in Mathematics No. 203, Springer-Verlag, New York, 1971.

ACQUAINTANCE GRAPH PARTY PROBLEM

A. J. SCHWENK, University of Michigan

The purpose of this note is to present a new, short solution to a problem of A. W. Goodman and to add a related result. Goodman [2] posed and solved the following generalization of an elementary problem from the MONTHLY [1]:

A party yields a graph G if we let the people at the party be represented by points and then let two points be adjacent if those two people are acquainted. A **full triangle** in G corresponds to a subset of three people who are mutually acquainted. An **empty triangle** in G corresponds to a subset of three people who are mutually strangers. Goodman found:

THEOREM 1. *Let E and F be the number of empty and full triangles respectively, and let square brackets denote the usual greatest integer function. Then in every graph with p points*

$$E + F \geqq \binom{p}{3} - \left[\frac{p}{2}\left[\left(\frac{p-1}{2}\right)^2\right]\right] \tag{1}$$

and this lower bound is sharp for each positive integer p.

Proof. Let P be the number of **partial triangles** in G, that is, the number of triangles containing exactly one or two lines. It is evident that

$$E + F + P = \binom{p}{3}. \tag{2}$$

We shall follow the notation in [3]. In particular, let d_i be the **degree** of the point v_i, i.e., the number of people acquainted with the ith person. For each point v_i, every choice of a pair of points consisting of one of his d_i acquaintances and one of his $p - 1 - d_i$ nonacquaintances produces a partial triangle. Thus, each v_i produces $d_i(p - 1 - d_i)$ partial triangles. Furthermore, we note that every partial triangle is counted twice in this manner (once for each endpoint in the triangle). Consequently,

$$P = \frac{1}{2} \sum_{i=1}^{p} d_i(p - 1 - d_i). \tag{3}$$

In view of equation (2), we minimize $E + F$ by maximizing P.

Each term of the sum is a quadratic function of d_i. If p is odd, the maximum value of $(p-1)^2/4$ is attained for each term when $d_i = (p-1)/2$. If p is even, the maximum permitted value of $p(p-2)/4$ is attained when $d_i = p/2$ or $(p-2)/2$. In either case, we note that this maximum value can be expressed as

$$\left[\left(\frac{p-1}{2}\right)^2\right],$$

and so

$$(4)\qquad P \leqq \frac{1}{2}\sum_{i=1}^{p}\left[\left(\frac{p-1}{2}\right)^2\right] = \frac{p}{2}\left[\left(\frac{p-1}{2}\right)^2\right].$$

But, since P is an integer, we may strengthen this to read

$$(5)\qquad P \leqq \left[\frac{p}{2}\left[\left(\frac{p-1}{2}\right)^2\right]\right].$$

Equations (2) and (5) now yield the desired bound:

$$(6)\qquad E+F = \binom{p}{3} - P \geqq \binom{p}{3} - \left[\frac{p}{2}\left[\left(\frac{p-1}{2}\right)^2\right]\right].$$

Next, for each p, we must find a graph G_p attaining this bound. But equality in equation (1) is equivalent to equality in equation (5) which occurs only when

$$(7)\qquad P = \frac{1}{2}\sum_{i=1}^{p} d_i(p-1-d_i) = \left[\frac{p}{2}\left[\left(\frac{p-1}{2}\right)^2\right]\right].$$

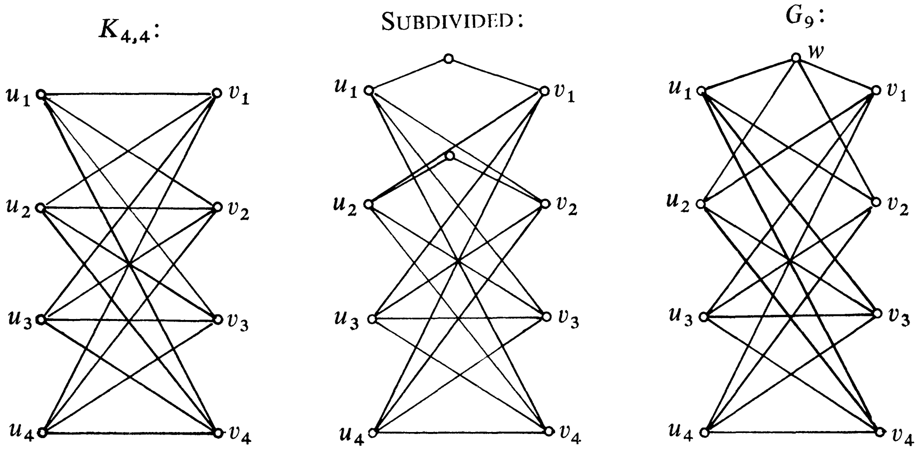


FIG. 1. The construction of G_9.

If $p=2n$, let G_p be the complete bipartite graph $K_{n,n}$. Now G_p is regular of degree n, and is seen to satisfy equation (7). If $p = 2n + 1$, the construction of G_p is a bit more involved. As depicted for $p = 9$ in Figure 1, we start with $K_{n,n}$ with its points labeled $u_1, u_2, \cdots, u_n$; $v_1, v_2, \cdots, v_n$ and we subdivide line u_iv_i for $i \leqq n/2$. We now obtain G_p by identifying these $[n/2]$ subdivision points to form a single point labeled w. We observe that G_p has $2n$ points of degree n and one point of degree $2[n/2]$. It is routine to check that equation (7) is satisfied, completing the proof.

L. Sauvé [5] also gave a shortened proof of Goodman's result. He then noted that when p is even, $E+F$ can be minimized while keeping $F=0$, but when p is odd and greater than seven, a proof due to P. Erdös (which Sauvé presented) demonstrates that $F>0$ for all graphs attaining the minimum for $E+F$. We now refine this result.

THEOREM 2. *In every graph attaining the minimum possible value for* $E+F$,

$$F \geqq \begin{cases} 0 & \text{if } p=2n \\ n(n-1) & \text{if } p=4n+1 \text{ or } 4n+3 \end{cases} \tag{8}$$

and this lower bound is sharp for each positive integer p.

Proof. We first observe that this bound is attained by the graphs G_p constructed above. If $p=2n$, then $G_p=K_{n,n}$ which obviously has no full triangles. If $p=4n+1$ or $4n+3$, we notice that G_p-w is a bigraph, and so, has no full triangles. Thus, every full triangle of G_p has the form $u_i v_j w$. But, recalling the construction of G_p, we see that this will be a full triangle if and only if $i \leqq n$, $j \leqq n$, and $i \neq j$. Consequently, G_p has $F=n(n-1)$ as desired.

It remains to be shown that the bound in equation (8) can not be violated. When p is even, the inequality is trivial, so we need only consider what happens when p is odd. When $n=1$, the bound is again trivial, so we may assume $n \geqq 2$.

CASE 1. $p=4n+1$. From among the graphs minimizing $E+F$, select H to be one with the smallest value of F. By the above discussion, $F \leqq n(n-1)$. A point of H lies in an average of $3F/(4n+1)$ full triangles. Thus, there exists a point v_0 lying in t full triangles where

$$t \leqq \frac{3F}{4n+1} \leqq \frac{3n(n-1)}{4n+1} < n-1. \tag{9}$$

Since t and n are integers, we may strengthen this to read $t \leqq n-2$. Let V be the point set of H, let $A=\{v \in V \mid v \text{ is adjacent to } v_0\}$, and let $B=V-(A \cup \{v_0\})$.

In order to minimize $E+F$, graph H must satisfy equation (7) by being regular of degree $2n$. Consequently, $|A|=2n=|B|$. Now since v_0 lies in t full triangles, there are exactly t lines whose endpoints both lie in set A. A simple count now reveals that there must be $4n^2-2n-2t$ lines joining sets A and B in order to contribute enough lines so that each point of A has degree $2n$. Finally, in order to fill out the degrees of the points in set B, there must be $n+t$ lines within set B.

Consider a line x in set B. Its two endpoints are incident with $4n-2$ other lines, of which at most $n+t-1$ can lie in set B. Thus, at least $3n-t-1$ of them have an endpoint in A. Since $|A|=2n$ and no point of A can lie on more than two of these lines, we conclude that at least $(3n-t-1)-2n=n-t-1$ points of set A lie on exactly two of these lines. Each such point determines a full triangle containing x.

Thus, each line of B lies in at least $n - t - 1$ full ABB triangles, and H has a total of at least $(n + t)\ (n - t - 1)$ full ABB triangles.

Similarly, consider a line y in set A. It is seen to be incident with $4n - 2$ other lines, of which two have v_0 as an endpoint and at most $t - 1$ lie in set A. Thus, at least $4n - t - 3$ of the lines have an endpoint in set B. Consequently, at least $2n - t - 3$ points of set B lie on exactly two of these lines. Each such point determines a full AAB triangle containing y. Thus, each line of set A lies in at least $2n - t - 3$ full AAB triangles, and H has at least $t(2n - t - 3)$ full AAB triangles.

Recalling the t full triangles containing v_0, we may add t to the sum of the two bounds found above to conclude that

$$F \geqq t + (n + t)(n - t - 1) + t(2n - t - 3) = n^2 - n + 2t(n - t - \tfrac{3}{2}).$$

Finally, since $0 \leqq t \leqq n - 2$, we conclude that $F \geqq n^2 - n$ as desired.

CASE 2. $p = 4n + 3$. As in Case 1, we select H to have minimal F, and we observe that $F \leqq n(n - 1)$. In order to minimize $E + F$, this graph must satisfy equation (7) by having $p - 1$ points of degree $2n + 1$ and one point w of degree $2n$ or $2n + 2$. (Note that a graph with an odd number of points cannot be regular with odd degree, [**3**, Cor. 2.1 (a)].) We may as well assume w has degree $2n$, for if it were $2n + 2$, we could delete a line incident with w producing a new graph H' with the desired degrees and $F' \leqq F$ full triangles.

As in Case 1, we wish to select a point v_0 which doesn't lie in too many full triangles. However, the possibility $v_0 = w$ is troublesome. Consequently, we select v_0 to be a point of degree $2n + 1$ lying in the smallest number of full triangles, say t. Then, regardless of how few triangles w lies in, we can at least be sure that $t(4n + 2) \leqq 3F \leqq 3n(n - 1)$. Thus

$$t \leqq \frac{3n(n-1)}{4n+2} < n - 1. \tag{10}$$

So, as in Case 1, $t \leqq n - 2$.

We now define sets A and B as in Case 1 and carry out the same counting procedure, only now we must consider two subcases depending upon the location of w.

Subcase (i). $w \in A$. Proceeding as in Case 1, we find $n + t + 1$ lines in set B each lying in at least $n - t - 1$ full ABB triangles. We see that each of the t lines in set A lies in at least $2n - t - 3$ full AAB triangles. Adding the t full triangles containing v_0 we have a bound of

$$F \geqq t + (n + t + 1)(n - t - 1) + t(2n - t - 3) = n^2 - 1 + 2t(n - t - 2).$$

Since $0 \leqq t \leqq n - 2$, we have $F \geqq n^2 - n$ as required.

Subcase (ii). $w \in B$. In this subcase, we find set B has $n + t$ lines each lying in at least $n - t - 1$ full ABB triangles and each of the t lines in set A lies in at least $2n - t - 2$ full AAB triangles. Adding the t full triangles containing v_0 we have a bound of,

$$(11) \qquad F \geqq t + (n + t)(n - t - 1) + t(2n - t - 2) = n^2 - n + 2t(n - t - 1).$$

Since $0 \leqq t \leqq n - 2$, this yields $F \geqq n^2 - n$, completing the proof.

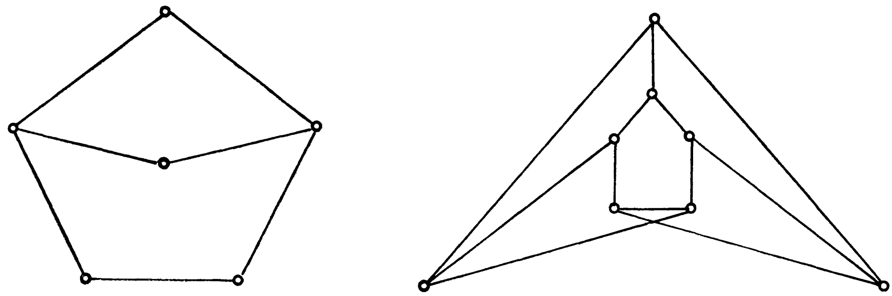


FIG. 2. Two graphs minimizing $E + F$ while keeping $F = 0$.

In conclusion, we observe that when p is odd G_p is the unique graph attaining the bound in Theorem 2. The proof of this fact is straightforward, but too cumbersome to merit inclusion here. On the other hand, when $p = 2n$, many graphs attain the bound. For example, those subgraphs obtained by removing a set of independent lines from $K_{n,n}$ do so, and certain other graphs also work. Two of these graphs are shown in Figure 2 for $p = 6$ and 8.

Acknowledgments. The author is grateful to P. Erdös and to F. Harary for their helpful comments.

References

1. C. W. Bostwick, E1321, this MONTHLY, 65 (1958) 446; 66 (1959) 141–142.
2. A. W. Goodman, On sets of acquaintances and strangers at any party, this MONTHLY, 66 (1959) 778–783.
3. F. Harary, Graph Theory, Addison-Wesley, Reading, Mass. 1969.
4. G. Lorden, Blue-empty chromatic graphs, this MONTHLY, 69 (1962) 114–120.
5. L. Sauvé, On chromatic graphs, this MONTHLY, 68 (1961) 107–111.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

PROBLEMS ON THE DENSITY OF ARITHMETIC SEQUENCES

A. A. MULLIN, Arlington, Virginia,

Dedicated to the memory of Professor Hans Rademacher

Apply the Unique Factorization Theorem to itself, inductively; i.e., if the natural number n has the standard form

$$n = p_1^{k_1} \cdots p_m^{k_m}$$

apply the Unique Factorization Theorem to each and every k_j, and repeat this process with successively generated exponents until a unique "constellation" of prime numbers *alone* is obtained, called a **mosaic**. E.g., the mosaic of

$$10{,}000 \text{ is } 2^{2^2} 5^{2^2}.$$

Clearly, if the mosaics of a set of integers have no prime number in common, then those integers are relatively prime, but not conversely, in general. Using this result, it is relatively easy to extend or modify classical number-theoretic concepts whose definitions use the notion of the greatest common divisor [**1** and **2**].

By analogy to basic density-theoretic results of P. Erdös [**3**], an integer n is said to have the **property** M if every finite sequence of consecutive integers which contains n also contains an integer whose mosaic has no prime in common with the mosaics of all the other integers in the sequence. It can be quickly shown that every prime number enjoys the property M. On the other hand, infinitely many integers do *not* enjoy the property M; e.g., p^2 does *not* enjoy the property M for any odd prime p. Indeed, it can be shown that the density ([**4**], p. xix) of integers *not* enjoying the property M is strictly positive. Determine whether or not the lower density d (i.e., $d = \liminf_{n\to\infty} (A(n)/n)$, where $A(n)$ is the number of members in the sequence A not exceeding n) of integers with the property M is strictly positive. If so, one has as an immediate corollary that the analogous lower density determined by Erdös [**3**] is strictly positive, too.

References

1. A. A. Mullin, On Möbius function and related matters, this MONTHLY, 74 (1967) 1100–1102.
2. ———, Additivity and multiplicativity, Solution to Problem 5248 (1964, 1138), this MONTHLY, 72(1965) 1140–1141.
3. P. Erdös, Some remarks on number theory, Israel J. Math., 3 (1965) 6–12.
4. H. Halberstam and K. F. Roth, Sequences (vol. 1), Oxford Univ. Press, 1966.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

DECOMPOSING MODULES OVER A PRINCIPAL IDEAL DOMAIN

R. P. HOLTEN, Cheyney State College, Pennsylvania

A standard and elegant way to obtain the Fundamental Theorem for torsion abelian groups and the rational or Jordan canonical form theorem for matrices is to view both results as results on the decomposition of torsion modules over a principal ideal domain (P.I.D.) (see, for example, [**4**, chapter XV] or [**2**, Chapter XV]). In this note we offer a new perspective of these results.

Throughout, R will denote a P.I.D. An R-module V is *torsion* if for each x in G there is an $r \neq 0$ in R such that $rx = 0$. Set $\operatorname{ord}(x)$ = a generator of the ideal consisting of those s in R with $sx = 0$.

Examples of torsion modules: A finite abelian group is a torsion R-module, where R is the ordinary integers. If V is a vector space over a field K and T is a linear operator on V whose image is finite-dimensional, then V is a torsion $K[T]$-module.

Our approach is to describe a useful condition on a submodule A of V so that A is a direct summand of V, that is, there exists another submodule B with $V = A \oplus B$. The condition we describe is motivated by the well-known fact that an abelian group A is a direct summand of each abelian group containing it if and only if A is divisible—that is, the equation $rx = a$ is solvable in A for any a in A, $r \neq 0$ any integer ([**3**], p. 93, [**1**], p. 8).

DEFINITION: Let R be a P.I.D. and let V be an R-module. A submodule A of V is called **V-divisible** if whenever the equation $rx = a + b$ is solvable in V for r in R, a in A, and b in V such that $A \cap Rb = \{0\}$, then the equation $rx = a$ is solvable in A.

A related concept is that of purity: a submodule A of V is pure [**1**, p. 14] if whenever the equation $rx = a$, $r \in R$, $a \in A$, is solvable in V, it is solvable in A. For a submodule A of V, A is divisible $\Rightarrow$ A is V-divisible $\Rightarrow$ A is a direct summand of

$V \Rightarrow A$ is a pure submodule of V. None of the reverse implications is true: for the first, take $V = A$; for the second, take $R = A = Z$, the usual integers, $V = Z \oplus Z$, $a = (3,0)$, $b = (3,12)$, $r = 2$. For the last see [**1**, p. 14 (i)].

We prove the second implication as

THEOREM 1. *Let R be a P.I.D. and V an R-module. If A is a V-divisible R-submodule of V, then A is a direct summand of V.*

The proof is almost identical to the proof of the corresponding result with A divisible [**1**, Theorem 2, p. 8]:

Proof. Let B be a maximal submodule of V such that $A \cap B = \{0\}$, and let v be any non-zero element of V. If v is not in B, then $B + Rv$ intersects A non-trivially, so there exists r in R and b' in B with $0 \neq b' + rv$ in A. So rv is in $A \oplus B$. Since R is a P.I.D., r factors into a product of a finite number of primes. Let $r = pr'$ with p prime. We shall show that $r'v$ is in $A \oplus B$, whence by a finite number of repetitions of the argument it will follow that v is in $A \oplus B$. Set $g = r'v$.

We have $pg = a + b$, a in A, b in B. Using V-divisibility, find an a' in A so that $pa \ = a$. Set $g' = g - a'$. Then $pg' = b$ is in B. Either g' is in B, in which case we are done, or else there are b'' in B, s in R with $0 \neq b'' + sg'$ in A. If p were a factor of s, then sg' would be in B and $A \cap B \neq \{0\}$. So p and s are relatively prime. Find t, u in R such that $tp + us = 1$. Then

$$g' + ub'' = u(b'' + sg') + tpg'$$

is in $A \oplus B$, so that g', and hence g, is in $A \oplus B$. That completes the proof.

We now use the notion of V-divisibility to obtain the standard decomposition theorems for torsion modules over a P.I.D.

The primary decomposition theorem for a torsion R-module V says that $V_p = \{x \text{ in } V \mid p^k x = 0 \text{ for some } k\}$, the p-primary component of V, is a direct summand of V for each prime p of R, and V is the direct sum of the V_p's.

THEOREM 2. *Let V be a torsion R-module; then V_p is a direct summand of V.*

Proof. By Theorem 1 it suffices to show that V_p is V-divisible. Let g be a solution of $rx = a + b$, where a is in)pV_p and or$b$) $= s$ with p and s relatively prime.

Let $r = r_1 p^k$ with p and r_1 relatively prime. Since $rsg = sa$ is in V_p we have $r_1 sg$ in V_p also. Put $p^m = \text{ord}(a)$ and find t_1, t_2, t_3, t_4 so that

$$t_1 r_1 s - t_2 p^m = 1 \quad \text{and} \quad t_3 s - t_4 p^m = 1 .$$

Finally, $a^1 = t_3 s t_1 r_1 sg$ is in V_p and is a solution of $rx = a$ because

$$ra^1 = t_1 r_1 s t_3 rsg = t_1 r_1 s t_3 sa = t_1 r_1 s(1 + t_4 p^m)a = t_1 r_1 sa = (1 + t_2 p^m)a = a .$$

Since V_p is V-divisible it now follows easily that the direct sum of any family of distinct p-primary components is again V-divisible and hence V is the direct sum of the V_p's.

The invariant factors theorem for a finitely generated torsion R-module V which is p-primary (that is, $V = V_p$) says that V_p is a direct sum of cyclic submodules. This result also follows easily from Theorem 1.

THEOREM 3. *Let $V = V_p$ be a p-primary R-module such that $p^n V = \{0\}$ for some n. Then if a is an element of order p^N, N maximal, Ra is a direct summand of V.*

Proof. Let a be an element of maximum order p^N in V and let $A = Ra$. Let g be a solution of $rx = sa + b$ in V, with $A \cap Rb = \{0\}$. Put $r = r_1 p^m$ and $s = s_1 p^n$, with r_1 and s_1 relatively prime to p. Now

$$0 = rp^{N-m}g = s_1 p^{n+N-m}a + p^{N-m}b.$$

Since $A \cap Rb = \{0\}$, it follows that $s_1 p^{n+N-m}a = 0$, so $n \geqq m$ and the ideal generated by p^N and r is generated by p^m, where p^m divides s. Solve $zr + wp^N = s$ and put $a' = za$ in A. Then $ra' = sa - wp^N a = sa$. So $Ra = A$ is V-divisible and the result follows from Theorem 1.

Prufer's theorem for V_p can be obtained easily by using induction and a proof similar to the proof of Theorem 3.

Acknowledgment: The author wishes to thank Professor Lindsay Childs for revising the format and adding a number of expository remarks to the paper.

References

1. I. Kaplansky, Infinite Abelian Groups, rev. ed. Univ. of Michigan, Ann Arbor, 1969.
2. S. Lang, Algebra, Addison-Wesley, Reading, Mass., 1965.
3. S. MacLane, Homology, Academic Press, New York, 1963.
4. B. L. van der Waerden, Modern Algebra, Ungar, New York, 1950.

EVERY CONVEX FUNCTION IS LOCALLY LIPSCHITZ

Wayne State University, Mathematics Department Coffee Room

A real-valued function f defined on a convex subset Ω of $\mathbb{R}^n$ is said to be **convex** if

$$f(tx + (1-t)y) \leqq tf(x) + (1-t)f(y), \tag{1}$$

whenever x and y are in Ω and $0 \leqq t \leqq 1$. A real-valued function f defined on an arbitrary subset K of $\mathbb{R}^n$ is said to be **Lipschitz on K with Lipschitz constant M if**

$$|f(x) - f(y)| \leqq M \| x - y \| \tag{2}$$

for every x and y in K (in this paper, $\| x - y \|$ will always denote the distance in $\mathbb{R}^n$ between x and y). The purpose of this article is to give what the authors feel is an interesting proof of the following:

THEOREM. *If f is a real-valued function defined on a convex open subset Ω of $\mathbb{R}^n$, if f is convex, and if $K \subset \Omega$ is compact, then there is a Lipschitz constant M such that* (2) *above holds for every $\mathbf{x}$ and $\mathbf{y}$ in K.*

COROLLARY. *Let f be a convex function defined on a convex open subset Ω of $\mathbb{R}^n$. Then f is continuous on Ω.*

Proof of Corollary. Let $\mathbf{x}$ be any point of Ω. Then there is a closed ball B centered at $\mathbf{x}$ with radius $r > 0$ such that $B \subset \Omega$. Since B is compact the theorem above applies, so that f is Lipschitz on B with some Lipschitz constant M. Given any $\varepsilon > 0$ we may choose δ to be the smaller of the two numbers r and ε/M so that

$$\|\mathbf{y} - \mathbf{x}\| < \delta \text{ implies } |f(\mathbf{y}) - f(\mathbf{x})| < \varepsilon.$$

LEMMA. *If $\alpha < \beta < \gamma$ are points of an open interval $\Omega \subseteq \mathbb{R}^1$ and if f is convex on that interval, then*

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} \leqq \frac{f(\gamma) - f(\alpha)}{\gamma - \alpha} \leqq \frac{f(\gamma) - f(\beta)}{\gamma - \beta}.$$

Proof of Lemma. Choose t so that $0 < t < 1$ and $t\alpha + (1-t)\gamma = \beta$. Then by (1),

$$f(\beta) = f(t\alpha + (1-t)\gamma) \leqq tf(\alpha) + (1-t)f(\gamma).$$

Solving for t and verifying that the inequality above implies the inequalities of the lemma involves only straightforward algebraic manipulation.

Proof of Theorem. The proof is by induction on the dimension n of the space $\mathbb{R}^n$ containing Ω as a subset. Suppose first that $n = 1$, so that Ω is an open interval of $\mathbb{R}^1$. Let K be any compact subset of Ω. We can then choose points $a < b < c < d$ in Ω such that for any x and y in K with $x < y$ we have $a < b < x < y < c < d$. Thus by repeated applications of the above lemma,

$$\frac{f(b) - f(a)}{b - a} \leqq \frac{f(y) - f(x)}{y - x} \leqq \frac{f(d) - f(c)}{d - c}. \tag{3}$$

This proves (2) for the case $n = 1$, since M can be taken to be the larger of the two numbers

$$\left|\frac{f(b) - f(a)}{b - a}\right| \text{ and } \left|\frac{f(d) - f(c)}{d - c}\right|.$$

Now assume that the theorem is true for $n = k - 1$ ($k \geqq 2$), let f be a convex function defined on a convex open subset Ω of $\mathbb{R}^k$, and let K be a compact subset of Ω. We are going to find compact sets X and Y such that

(a) $K \subseteq X \subseteq Y \subseteq \Omega$

(b) K, ∂X, and ∂Y are pairwise disjoint

(c) both X and Y are finite unions of k-dimensional boxes with edges parallel to the coordinate axes.

Suppose for the moment that such sets X and Y have been found.

The sets ∂X and ∂Y are compact and are the union of certain $(k-1)$-dimensional "faces" of "boxes". Let H be the coordinate hyperplane which intersects a "box" in such a "face". The function f restricted to $H \cap \Omega$ can be considered as a convex function on a convex open subset of $\mathbb{R}^{k-1}$, and is therefore continuous on this set by the induction assumption and the corollary above. Thus f is continuous on ∂X and ∂Y, so that the function

$$Q(x,y) = \frac{|f(x) - f(y)|}{\|x - y\|} \qquad (x \in \partial X,\ y \in \partial Y)$$

is continuous (since ∂X and ∂Y are compact and disjoint, the denominator is bounded away from zero). Since $(\partial X) \times (\partial Y)$ is compact, Q takes a maximum value M on this set.

Now consider any $\boldsymbol{x}$ and $\boldsymbol{y}$ in K with $\boldsymbol{x} \neq \boldsymbol{y}$. The unique line ℓ through $\boldsymbol{x}$ and $\boldsymbol{y}$ must cut ∂X in points $\boldsymbol{b}$ and $\boldsymbol{c}$ and must cut ∂Y in points $\boldsymbol{a}$ and $\boldsymbol{d}$ such that the order of these points on ℓ is $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}$ ($\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$, and $\boldsymbol{d}$ are not necessarily unique); indeed, starting at $\boldsymbol{y}$ and traveling on ℓ in the direction away from $\boldsymbol{x}$ we must leave the set X (encountering $\boldsymbol{c} \in \partial X$) and later must leave Y (encountering $\boldsymbol{d} \in \partial Y$) — $\boldsymbol{a}$ and $\boldsymbol{b}$ may be found similarly.

We may consider f restricted to $\ell \cap \Omega$ as a convex function on an open interval of $\mathbb{R}^1$ and apply the result (3) to obtain

$$\frac{|f(\boldsymbol{y}) - f(\boldsymbol{x})|}{\|\boldsymbol{y} - \boldsymbol{x}\|} \leqq \max\left[\frac{|f(\boldsymbol{d}) - f(\boldsymbol{c})|}{\|\boldsymbol{d} - \boldsymbol{c}\|}, \frac{|f(\boldsymbol{b}) - f(\boldsymbol{a})|}{\|\boldsymbol{b} - \boldsymbol{a}\|}\right] \leqq M.$$

This proves that (2) holds for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in K.

It remains to find X and Y. Since K is compact and Ω is open, there is an $r > 0$ such that any point of $\mathbb{R}^k$ must be in Ω if it is closer than r to K. For any k-tuple $(m_1, \cdots, m_k)$ of integers, define the "box"

$$B(m_1, \cdots, m_k) = \left\{(x_1, \cdots, x_k) \in \mathbb{R}^k;\ \frac{m_i r}{10\sqrt{k}} \leqq x_i \leqq \frac{(m_i + 1)r}{10\sqrt{k}} \text{ for } i = 1, \cdots, k\right\}.$$

These "boxes" cover $\mathbb{R}^k$. The length of each side is $r/(10\sqrt{k})$ and the length of each main diagonal is $r/10$.

Let X be the union of all those "boxes" which are closer than $r/5$ to K. Let Y be the union of all those "boxes" which are closer than $4r/5$ to K. Let ∂X and ∂Y denote the boundaries of these two sets. *The sets ∂X, ∂Y and K are pairwise disjoint*; for example, if $\boldsymbol{x} \in K$ then by the triangle inequality any "box" adjacent to a "box" containing $\boldsymbol{x}$ is contained in both X and Y, so that $\boldsymbol{x} \notin \partial X$ and $\boldsymbol{x} \notin \partial Y$. This completes the proof of the theorem.

REMARK. The proof of the theorem in the case $n = 1$ follows closely exercise 17.37(a) and (b) of Hewitt and Stromberg [**1**], pp. 271–272.

Reference

1 E. Hewitt, and K. Stromberg, Real and abstract analysis, Springer-Verlag, New York, 1965.

THE DERIVATIVE OF A DETERMINANT

M. A. GOLBERG, University of Nevada

1. Introduction. In courses on differential equations one of the basic formulas which is developed is Wronski's expression for the determinant of the fundamental matrix of a linear differential equation. The formula is usually derived by setting up a scalar equation for the derivative of the determinant. In this note we present a convenient representation for the derivative of an arbitrary determinant valued function which has as immediate consequence Wronski's formula. This derivation seems to be more straightforward than the ones usually presented [**2**].

2. Notation. For a given $n \times n$ matrix $A = \{a_{ij}\}$, $(i,j) = 1,2,\cdots,n$, tr A will denote its trace, Adj A its classical adjoint, and det A its determinant. Cof a_{ij} will denote the cofactor of the ijth element of A. Derivatives will be denoted by subscripts.

3. Main Theorem. *Let U be an open subset of the reals. Let $A(t)$ be a differentiable matrix valued function on U. Then $d(t) \equiv \det A(t)$ is differentiable and satisfies*

$$d_t(t) = \operatorname{tr}((\operatorname{Adj} A(t))A_t(t)),\ t \in U. \tag{1}$$

Proof. The differentiability of $d(t)$ is standard [**2**]. Let

$$A(t) = \{\alpha_1(t),\alpha_2(t),\cdots,\alpha_n(t)\}, \tag{2}$$

where $\alpha_i(t)$ is the ith row of $A(t)$. Then

$$d_t(t) = \sum_{i=1}^{n} \det\{\alpha_1(t),\alpha_2(t),\cdots,\alpha_{i,t}(t),\cdots,\alpha_n(t)\}. \tag{3}$$

Now

$$\operatorname{Adj} A(t) = \{\operatorname{cof} a_{ji}(t)\},\ (i,j = 1,2,\cdots,n), \tag{4}$$

so that

$$(\operatorname{Adj}(A(t)))A_t(t) = \left\{\sum_{k=1}^{n} \operatorname{cof} a_{ki}(t) a_{kj,t}(t)\right\},\ (i,j = 1,2,\cdots,n). \tag{5}$$

Therefore

$$\text{(6)} \qquad \operatorname{tr}((\operatorname{Adj} A(t))A_t(t)) = \sum_{i=1}^{n} \left\{ \sum_{k=1}^{n} \operatorname{cof} a_{ki}(t) a_{ki,t}(t) \right\}$$
$$= \sum_{k=1}^{n} \left\{ \sum_{i=1}^{n} \operatorname{cof} a_{ki}(t) a_{ki,t}(t) \right\}.$$

Using the Laplace expansion of a determinant [**1**], we see that

$$\text{(7)} \qquad \sum_{i=1}^{n} \operatorname{cof} a_{ki}(t) a_{ki,t}(t) = \det\{\alpha_1(t), \alpha_2(t), \cdots, \alpha_{k,t}(t), \cdots, \alpha_n(t)\}.$$

Therefore

$$\text{(8)} \qquad \operatorname{tr}((\operatorname{Adj} A(t))A_t(t)) = \sum_{k=1}^{n} \det\{\alpha_1(t), \alpha_2(t), \cdots, \alpha_{k,t}(t), \cdots, \alpha_n(t)\}$$
$$= d_t(t).$$

COROLLARY. *If for each $t \in U$, $A(t)$ is invertible, then*

$$\text{(9)} \qquad d_t(t) = [\operatorname{tr}(A^{-1}(t)A_t(t))]d(t).$$

Proof. By Cramer's Rule, $\operatorname{Adj} A(t) = d(t)A^{-1}(t)$, $t \in U$. So by the theorem

$$d_t(t) = \operatorname{tr}(d(t)A^{-1}(t)A_t(t)) = [\operatorname{tr}(A^{-1}(t)A_t(t))]d(t).$$

COROLLARY. (Wronski's Formula [**2**].) *Let $A(t)$ satisfy the matrix differential equation*

$$\text{(10)} \qquad A_t(t) = B(t)A(t), \qquad t \in U.$$

where $B(t)$: $U \to \operatorname{Hom}(R^n)$ is continuous. Then $d(t)$ satisfies the scalar equation

$$\text{(11)} \qquad d_t(t) = [\operatorname{tr} B(t)]d(t).$$

Proof. By the theorem and (10),

$$d_t(t) = \operatorname{tr}((\operatorname{Adj} A(t))B(t)A(t))$$
$$= \operatorname{tr}(A(t)\operatorname{Adj} A(t)B(t)),$$

by the symmetry of trace. So using Cramer's rule again we get that

$$\text{(12)} \qquad d_t(t) = \operatorname{tr}(d(t)B(t)) = [\operatorname{tr} B(t)]d(t).$$

H. Hochstadt, in a previous paper [**3**] in this journal has also given several proofs of the corollary to the main theorem above. His, and as far as the author is aware, all other proofs make specific use of properties of the Wronskian. In this paper equation (11) is derived as a consequence of a formula for the derivative of an arbitrary determinant valued function. The fact that the derivative of the Wronskian determinant is obtainable in this way appears to have been overlooked. The proof presented here appears to be simpler and more straightforward than previous proofs of this result.

References

1. Daniel T. Finkbeiner II, Introduction to Matrices and Linear Transformations, Freeman, San Francisco, 1966.
2. Einar Hille, Lectures on Ordinary Differential Equations, Addison-Wesley, Reading, Mass., 1969.
3. H. Hochstadt, On the Derivative of the Wronskian Determinant, this MONTHLY, 75 (1968) 767–772.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

A MODULAR APPROACH TO PREPARATORY MATHEMATICS

L. J. ABLON, Staten Island Community College

1. Introduction. On July 9, 1969, the Board of Higher Education of The City University of New York moved the target date for Open Admissions from 1975 to 1970. In their new policy the Board stated:

> We do not want to provide the illusion of an open door to higher education which in reality is only a revolving door, admitting everyone but leading to a high proportion of student failure after one semester.

The Board also provided support to

> insure that each unit of the University be given significant responsibilities for preparing the academically less-prepared student to engage in collegiate study.

This paper describes the response of the Mathematics Department of Staten Island Community College (SICC) to Open Admissions. The program described below is the fall 1971 form of a program first introduced in September 1970 for 216 students. There are now about 500 students in the program. The program is in a state of continuous evolution so the following describes what existed at the moment of writing and may or may not correspond to the situation at the moment of reading. No claim is made to originality in program structure or content. The program Proposal originally implemented in September 1970 under the terms of a contract with the Preparatory Skills Center of SICC was written by B. Greenberg, S. Richard, and M. Sormani, all from the Mathematics Department of SICC. It developed out

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before March 31, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2385. *Proposed by Burnett Meyer, University of Colorado*

Show that there exists no binary system with two-sided identity and two-sided inverses such that the associative law fails for exactly one ordered triple of elements.

E 2386. *Proposed by William Knight, University of New Brunswick*

The classical birthday problem can be phrased as a bet between a statistics teacher and a class of $n<365$ students, the teacher betting that at least two students have the same birthday. (The usual stake is one-up-ness rather than money.) If birthdays are (1) independently and (2) uniformly distributed over the 365 days of the year (leap years being ignored) the probability of the teacher's winning is $1-(365)_n/365^n$ where $(m)_n$ denotes the partial factorial $m!/(m-n)!$. But it is more likely that birthdays are not really equally numerous at all seasons. Show that this, in fact, makes the bet more favorable the the teacher; that is, if assumption (2) is dropped, $1-(365)_n/365^n$ is a lower bound attained only when all days are equally probable as birthdays.

E 2387. *Proposed by David Jacobson, Rutgers University*

It is well known that a Boolean ring with identity 1 is (von Neumann) regular

and 1 is the only unit in the ring. Conversely, show that if R is a commutative regular ring and 1 is the only unit in R, then R is a Boolean ring.

E 2388. *Proposed by A. W. Walker, Toronto, Canada*

Let $a, b, c; s, r, R, I, H$ denote the side lengths, semiperimeter, inradius, circumradius, incenter and orthocenter of a triangle ABC.

(i) For ABC arbitrary, prove that $bc + ca + ab \geqq (AI + BI + CI)^2$ with equality if and only if the triangle is equilateral.

(ii) For ABC non-obtuse, prove that $s^2 \geqq 2R^2 + 8Rr + 3r^2$ or, equivalently, $a^2 + b^2 + c^2 \geqq (AH + BH + CH)^2$, with equality if and only if the triangle is equilateral or right isosceles.

E 2389. *Proposed by Zbigniew Fiedorowicz, Illinois Institute of Technology*

Suppose that f is a strictly positive continuous function on the interval $[0, 1]$. Show that the following (two-sided) limit exists and find its value:

$$\lim_{\alpha \to 0} \left\{ \int_0^1 [f(x)]^\alpha dx \right\}^{1/\alpha}.$$

Can this result be generalized to a wider class of functions?

E 2390. *Proposed by Anon, Erewhon-upon-Yarkon*

Let $f(x)$ be continuous on (a, b) and suppose

$$D_s f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h} = 0$$

for all x in (a, b). Prove that $f(x)$ is constant.

SOLUTIONS OF ELEMENTARY PROBLEMS

Random Chords in an Ellipse

E 2324 [1971, 1020]. *Proposed by Frank Dapkus, Seton Hall University*

What is the probability that the length of a chord randomly drawn in an ellipse will not exceed the length of the minor axis? (By "randomly drawn chords" we mean those with midpoints uniformly distributed throughout the ellipse.)

Solution by Robert Patenaude, California Institute of Technology. Let the ellipse be given by $x^2/a^2 + y^2/b^2 = 1$ with $a \geqq b$. Laterally shrink the x-axis, changing it by a factor of b/a so that the ellipse becomes a circle. A chord of length $2b$ in the ellipse becomes a chord of length $2c(\theta)$ in the circle, where

$$c^2(\theta) = \frac{b^4}{a^2 \sin^2 \theta + b^2 \cos^2 \theta},$$

provided the distorted chord has θ as its angle of incidence to the y-axis. On the

ray elevated an angle θ from the x-axis the points which are centerpoints of chords in the circle exceeding $2c(\theta)$ in length are those points less than a distance of $r(\theta)$ from the origin, where $r^2(\theta) = b^2 - c^2(\theta)$. The complementary probability $1 - p$ can then be computed as the ratio of the area of these points to the area of the circle:

$$1 - p = \frac{1}{\pi b^2} \cdot 4 \int_0^{\pi/2} \tfrac{1}{2} r^2(\theta) d\theta = 1 - \frac{2b^2}{\pi} \int_0^{\pi/2} \frac{d\theta}{a^2 \sin^2\theta + b^2 \cos^2\theta}.$$

The latter definite integral is well known to be $\pi/2ab$, so that the desired probability is $p = b/a$.

Also solved by A. S. Adikesavan (India), Günter Bach (Germany), Robert Breusch, Jordi Dou (Spain), Harry Lass, R. W. W. Taylor, Charles Wexler, and the proposer.

Editor's comment. Three incorrect solutions were received. We note the similarity of this problem to "Bertrand's Paradox" which is that the probability that the length of a chord drawn randomly in a circle exceeds the length of the side of the inscribed equilateral triangle can be equal to 1/2, 1/3, or 1/4 depending on the argument used. (The paradox is resolved, of course by noting that each answer depends on a different idea of "random chord.")

A Binomial Coefficient Inequality

E2325 [1971, 1137]. *Proposed by S. I. Rosencrans, Tulane University*

Prove that if $-1 < \alpha < 0$

$$\binom{2\alpha}{2n} \geqq (2n+1)\binom{\alpha}{n}^2, \quad n = 0, 1, 2, \cdots,$$

while if $\alpha < -1$ the inequality is reversed.

Solution by R. L. Enison, Goddard Space Flight Center. Since $\binom{x}{0} = 1$ by definition, equality holds if $n = 0$. We will therefore assume $n \geqq 1$. Let $\beta = -\alpha$. A little computation shows that

$$\binom{2\alpha}{2n} \Big/ \binom{\alpha}{n}^2 = 2^n \binom{2n}{n}^{-1} \left(\frac{2\alpha - 1}{\alpha}\right)\left(\frac{2\alpha - 3}{\alpha - 1}\right) \cdots \left(\frac{2\alpha - 2n + 1}{\alpha - n + 1}\right)$$

$$= 2^n \binom{2n}{n}^{-1} \left(2 + \frac{1}{\beta}\right)\left(2 + \frac{1}{\beta + 1}\right) \cdots \left(2 + \frac{1}{\beta + n - 1}\right).$$

If $-1 < \alpha < 0$, then $0 < \beta < 1$ and thus

$$\binom{2\alpha}{2n} \Big/ \binom{\alpha}{n}^2 > 2^n \binom{2n}{n}^{-1} \left(\frac{3}{1}\right)\left(\frac{5}{2}\right) \cdots \left(\frac{2n+1}{n}\right) = 2n + 1.$$

Similarly if $\alpha < -1$, then $\beta > 1$ and the inequality is reversed. Evidently we have equality if and only if $n = 0$ or $\alpha = -1$.

Also solved by the proposer and thirty-four other readers.

A Generalized Ménage Problem

E 2326 [1971, 1137]. *Proposed by Harry Lass, California Institute of Technology*

Consider the following generalized ménage problem: N people labeled clockwise as $1, 2, 3, \cdots, N$, are seated at a table. If k people are chosen, labeled $1', 2', 3', \cdots, k'$, such that $1 \leqq 1' < 2' < 3' < \cdots < k' \leqq N$, we desire at least α_1 individuals between $1'$ and $2'$, at least α_2 individuals between $2'$ and $3'$, etc., and at least α_k individuals between k' and $1'$.

Show that the number of such choices is

$$\frac{k\alpha_k + N - \alpha}{N - \alpha}\binom{N-\alpha}{k} \quad \text{with} \quad \alpha = \sum_{i=1}^{k} \alpha_i.$$

Solution by M. T. Bird, San Jose State College. Let $f(n,r)$ be the number of ordered sequences of r nonnegative integers whose sum is n. We have

$$f(n,r) = \binom{n+r-1}{n}.$$

Let m_1 be the number of individuals between $1'$ and $2'$, m_2 the number of individuals between $2'$ and $3', \cdots$, and m_k the number of individuals between k' and $1'$. Let $n_i = m_i - \alpha_i$ and let n be the sum of the n_i. We have $n = N - k - \alpha$. Clearly each n_i may take on the values $0, 1, 2, \cdots, n$ subject to the limitation that their sum be n.

For a particular set $m_1, m_2, \cdots, m_k$ we may select the individual $1'$ to be the individual labeled $1, 2, \cdots, m_k + 1$ and hence we have $m_k + 1 = n_k + \alpha_k + 1$ choices for each set $m_1, m_2, \cdots, m_k$. For a particular set $\alpha_1, \alpha_2, \cdots, \alpha_k$ the number of choices that yield a particular m_k, (i.e., the number of choices of $n_1, n_2, \cdots, n_{k-1}$ such that their sum is $n - n_k$) is $f(n - n_k, k - 1)$. We conclude that the desired number of choices Q is

$$\begin{aligned} Q &= \sum_{j=0}^{n} (\alpha_k + j + 1) f(n-j, k-1) \\ &= \sum_{j=0}^{n} (\alpha_k + j + 1)\binom{n+k-j-2}{n-j} \\ &= \sum_{j=0}^{n} \binom{j+1}{1}\binom{n+k-j-2}{k-2} + \alpha_k \sum_{j=0}^{n} \binom{n+k-j-2}{k-2} \\ &= \binom{n+k}{k} + \alpha_k\binom{n+k-1}{k-1} = \binom{N-\alpha}{k} + \alpha_k\binom{N-\alpha-1}{k-1}. \end{aligned}$$

This reduces immediately to the form given in the statement of the problem.

Also solved by Arnold Adelberg, Jordi Dou (Spain), John Gaisser, M. G. Greening (Australia), J. C. Hickman, J. D. Hiscocks (Netherlands), David Kelly, Carolyn MacDonald, W. O. J. Moser, The Temple University Problem Solving Group, and the proposer.

Old Wine in New Bottles

E 2327 [1971, 1138]. *Proposed by Kenneth Rosen, University of Michigan*

Let S_m^n be the sum of the reciprocals of the integers not exceeding m and relatively prime to n. Prove that for $m > n$, $n \geqq 2$, S_m^n is never an integer.

Solution by H. S. Hahn, West Georgia College. Write

$$(*) \qquad S_m^n = \frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_s},$$

where $1 = a_1 < a_2 < \cdots < a_s$ are the integers not exceeding m which are relatively prime to n. It is obvious that a_2 is prime. Let a_2^k be the highest power of a_2 which does not exceed m. Then a_2^k is some a_i and is the only a_i which is divisible by a_2^k; this is because by choice of a_2, if $1 < c < a_2$, then $(c,n) > 1$ and if $c \geqq a_2$, then $ca_2^k \geqq a_2^{k+1} > m$.

Suppose to the contrary that S_m^n is an integer and let L denote the LCM of $a_1, a_2, \cdots, a_s$. (Note that $a_2^k \mid L$ but $a_2^{k+1} \nmid L$.) Multiply both sides of (*) by L and transfer all of the terms on the right to the left-hand side except the term L/a_2^k. Then the left-hand side is a multiple of a_2 but the right-hand side is not. This is a contradiction.

Also solved by Anders Bager (Denmark), Bennett College Team, R. E. Dressler, P. K. Garlick, Heiko Harborth (Germany), C. V. Heuer, C. V. Heuer & G. A. Heuer, Jeffrey Goodling, Erwin Just, O. P. Lossers (Netherlands), L. E. Mattics, Hugh Noland, C. B. A. Peck, St. Olaf College Students, T. Salát (Czechoslovakia), D. P. Sumner, Temple University Problem Solving Group, Charles Wexler, an anonymous solver, and the proposer.

Editor's comment: This problem is a variant of problems which have appeared previously. The granddaddy of them all is E 46 [1934, 48] which is (in the notation of the present problem) to show that S_m^1 is never an integer. Problem E 1964 [1967, 199] is to show that S_m^m is never an integer. (This was noted by several solvers.) Recently a proposal was received from Erwin Just and Norman Schaumberger which asks to show that the subsum of S_m^n taken over composite integers is not an integer (unless it is vacuous). It would seem that this area is pretty well exhausted.

Conditions which Make a Regular Semigroup a Group

E 2328 [1971, 1138]. *Proposed by D. R. Hayes, University of Massachusetts*

Suppose G is a semigroup having the property that, for every $a \in G$, there is a unique element $a^* \in G$ such that $aa^*a = a$. Prove that G is in fact a group.

Solution by D. E. Knuth, Stanford University. Since $a = aa^*a = aa^*aa^*a$ we must have $a^* = a^*aa^*$ for every $a \in G$. Now let $a, b \in G$ be arbitrary and let $x = a(ba)^*b$; then $xaa^*x = a(ba)^*baa^*a(ba)^*b = a(ba)^*ba(ba)^*b = a(ba)^*b = x$ and similarly $xb^*bx = x$ so that $aa^* = b^*b$. If we let e denote this common value and define $a^R = a^*$ we see that for every $a \in G, ae = aa^*a = a$ and $aa^R = e$.

But it is well known that a semigroup with a right identity and right inverses relative to this identity must be a group.

Also solved by the proposer and 56 other readers.

Editor's comment: Several solvers show that G must have a unique idempotent element and then use known results in semigroup theory to finish the problem. (See, e.g., Clifford and Preston, *Algebraic Theory of Semigroups*, Vol. I, AMS Colloquium Pub., p. 33.) D. M. Bloom mentions that this problem was assigned in an algebra class taught by Robert Taylor at Columbia University in 1955. The Temple University Problem Solving Group assert that a semigroup G with the property that for every $a \in G$ there exists a unique $\acute{a} \in G$ such that $a a \acute{a}$ is idempotent, must necessarily be a group. The multiplicative semigroup of 2×2 real matrices shows that the uniqueness assumption of the problem is vital.

Quite a few incorrect solutions were received. The most common error was for solvers to assume that since $a\,(a^*a) = a$, it must follow that a^*a is a right identity for G, forgetting that it is necessary to show that $ba^*a = b$ for every $b \in G$.

Two Functional Equations

E 2329 [1971, 1138]. *Proposed by R. S. Luthar, University of Wisconsin at Janesville*

Suppose that $0 < a < 1$ [so that $I = (0, a)$ is closed under multiplication].

(A) Find all continuous real-valued functions f defined on I which satisfy $f(xy) = xf(y) + yf(x)$.

(B) Find all continuous real-valued functions f defined on I which satisfy $f(xy) = xf(x) + yf(y)$.

Solution by Paul Chernoff, University of California, Berkeley. For part (A), let $g(x) = f(x)/x$. Then $g(xy) = g(x) + g(y)$, and it is known that this implies that $g(x) = C \log x$. That is, $f(x) = Cx \log x$ for some constant C.

For part (B), $f(x)$ must be identically zero, and there is no need to assume continuity *a priori*. Indeed, successively taking $y = x,\ x^2, x^3$ we get

$$\begin{aligned} f(x^2) &= 2xf(x) \\ f(x^3) &= xf(x) + x^2 f(x^2) = (x + 2x^3) f(x) \\ f(x^4) &= xf(x) + x^3 f(x^3) = (x + x^4 + 2x^6) f(x). \end{aligned}$$

However, $x^4 = x^2 \cdot x^2$, so that

$$f(x^4) = 2x^2 f(x^2) = 4x^3 f(x),$$

and hence $f(x) = 0$ except possibly when x is a root of the equation $2x^6 + x^4 + x = 4x^3$. That is, $f(x) = 0$ with at most a finite number of exceptions. But if $f(t) \neq 0$, then $f(t^2) \neq 0$, and this would lead to infinitely many points where $f(x) \neq 0$, a contradiction.

Also solved by Michael Barr, D. M. Bloom, R. L. Breisch, Frederick Carty, Neal Felsinger, John Gaisser, Michael Goldberg, M. G. Greening (Australia), Emil Grosswald, Ellen Hertz, G. A. Heuer, K. J. Heuvers, F. A. Homann, Marek Kuczma (Poland), Detlef Laugwitz (Germany), O. P. Lossers (Netherlands), Carolyn MacDonald, Beatriz Margolis (France), Bill Margolis, Oscar Ocelot (Israel), F. J. Papp, H. B. Potoczny, Jürg Rätz (Switzerland), Kenneth Rosen, St. Olaf College Students, St. Olaf Problem Group, P. S. Schnare, David Shelupsky, Susan B. Slesnick, Wolfe Snow, D. P. Sumner, Temple University Problem Solving Group, Charles Wexler, and A. C. Williams.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before March 31, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5884. *Proposed by Gérard Letac, University of Clermont, France*

Let $(X_j^{(i)})_{j=1}^{\infty}$, $i = 1, 2, \cdots, d$ be sequences of independent random variables with positive integer values and having distributions not depending on j. Denote

$$S_n^{(i)} = \sum_{j=1}^{\infty} X_j^{(i)}$$

and

$$S = \inf\{s\colon \text{there exist } n_1, \cdots, n_d \text{ such that } s = S_{n_i}^{(1)} = \cdots = S_{n_a}^{(d)}\}.$$

Prove that $E(X_1^{(i)}) < \infty$ for all $i = 1, 2, \cdots, d$ implies $S < \infty$ almost surely and $E(S) = E(X_1^{(1)}) \cdots E(X_1^{(d)})$.

5885. *Proposed by F. Haring and G. T. Nelson, North Dakota State University*

(a) Show:

$$1 + \frac{1}{2^2} - \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} + \frac{1}{8^2} - \frac{1}{10^2} - \frac{1}{11^2} + \cdots$$
$$= -\frac{2}{27}\pi^2 - 2\int_0^1 \frac{\ln x}{1 + x^3}\,dx.$$

(b) Find the sum of the series.

5886*. *Proposed by G. de Josselin de Jong, New Mexico Institute of Mining and Technology*

What is the maximal edge of a cube that can be placed inside a tesseract of edge 1?

5887. *Proposed by E. H. Umberger, Pennsylvania State University*

Find the radius ρ of the largest disk D for which there exists a continuous,

rectifiable curve C of unit length such that every point of D is within unit distance of some point of C.

5888*. *Proposed by Stanley Rajnak, Kalamazoo College*

Does there exist a real-valued function defined on $\mathbb{R}^2$ which has all partial derivatives of all orders at every point, but is not continuous on a dense set?

SOLUTIONS OF ADVANCED PROBLEMS

A Generalization of Fermat's Theorem

5807 [1971, 679]. *Proposed by E. G. Kundert, University of Massachusetts*

Put

$$\alpha(s; h, k) = \binom{s}{k}\binom{k}{s-h} = \frac{s!}{(s-h)!\,(s-k)!\,(h+k-s)!}$$

for $\max\{h, k\} \leqq s \leqq h + k$, and $\alpha(s; h, k) = 0$ otherwise. Let p be a fixed prime number and i, j fixed natural numbers. Prove

$$\Sigma\alpha(s_1; i, i)\alpha(s_2; i, s_1)\cdots\alpha(s_{p-2}; i, s_{p-3})\alpha(i+j; i, s_{p-2}) \equiv 0 \pmod{p},$$

where the summation is over all $s_1, s_2, \cdots, s_{p-2}$ such that

$$i \leqq s_1 \leqq s_2 \leqq \cdots \leqq s_{p-2} \leqq i + j.$$

Solution by the proposer. We refer to E. G. Kundert, *Structure theory in s-d-rings*, Nota I, Accademia dei Lincei (Rendiconti) Ser. VIII, Vol. XLI, fasc. 5, Nov. 1966, and let $x_0 = 1$, $x_i = s(x_{i-1})$, where $s(a)$ is defined as a mapping of an s-d-ring into itself given in the cited reference.

An induction using the methods of the reference yields

$$x_h \cdot x_k = \Sigma(-1)^{h+k+s}\alpha(s; h, k)x_s.$$

By induction on n,

$$dx_i^n = x_i^n - (x_i - x_{i-1})^n = \sum_{r=1}^{n} (-1)^{r-1}\binom{n}{r}x_i^{n-r}\cdot x_{i-1}^r.$$

Since $\binom{p}{r} \equiv 0 \pmod{p}$ for $1 \leqq r \leqq p - 1$, $dx_i^p \equiv x_{i-1}^p \pmod{p}$. Induction on i yields $x_i^p \equiv x_i \pmod{p}$ by "integrating" both sides of $dx_i^p \equiv x_{i-1}^p \equiv x_{i-1}$. If we write $x_i^p - x_i = \Sigma_{j\geqq 1}\beta_j x_{j+i} \equiv 0 \pmod{p}$, then $\beta_j \equiv 0 \pmod{p}$. But β_j is the left side of our formula if we disregard the sign. We used the fact that the operations d and s are preserved by the passage from the s-d-ring $\mathfrak{A}$ to the s-d-ring $\mathfrak{A}_p = \mathfrak{A}/(p)$ since (p) is deteal and inteal by #4 of the reference. $\mathfrak{A}_p$ is the s-d-ring over $\mathbb{Z}_p = \mathbb{Z}/(p)$.

We note that with $j = 1$ and $m = i + 1$, the formula becomes Fermat's theorem: $m^p - m \equiv 0 \pmod{p}$.

Editor's comment. The proposer's solution is the only one submitted. Elementary solutions would still be of interest.

Contractible Spaces

5809 [1971, 798]. *Proposed by Richard Stanley, Massachusetts Institute of Technology*

Let X be a topological space such that an arbitrary intersection of open sets is open. Show that if X is connected, compact and normal (without assuming T_0, T_1, or Hausdorff), then X is contractible.

Solution by Bill Beckmann, Davidson College. Given a point p in X, let U_p denote the intersection of all open sets containing p; by hypothesis, U_p is open. Since X is compact, there exists a minimal finite covering $U_1, \cdots, U_n$ of X, where $U_i = U_{p_i}$ for some point p_i in X. For each $i = 1, \cdots, n$, let V_i denote the union of all U_j, $j \neq i$. Suppose $U_i \cap U_k \neq \varnothing$ for $k \neq i$. Both $X - V_i$ and $X - V_k$ are closed subsets of X, and their intersection is empty. If p_i is not in $X - V_i$, then p_i lies in some U_j, $j \neq i$, and this contradicts the minimality of the covering. Therefore p_i is in $X - V_i$ and similarly p_k is in $X - V_k$. But this shows that if A and B are open sets containing $X - V_i$ and $X - V_k$ respectively, then $A \cap B \supset U_i \cap U_k \neq \varnothing$, contradicting the assumption of normality. Thus $U_i \cap U_k = \varnothing$. The connectedness of X implies that the minimal finite covering consists of one element, say U_1, and then $U_1 = X$.

Define $F: X \times I \to X$ by $F(x,t) = x$ if $0 \leqq t < 1$, and $F(x,1) = p_1$. Let G be any open set in X; if p_1 is not in G, then $F^{-1}(G) = G \times [0,1)$. And if p_1 is in G, then $G = X$. Hence F is the required homotopy.

Also solved by Michael Barr, Andreas Blass, E. N. Ferguson, D. A. Hejhal, A. A. Jagers (Netherlands), J. D. Klemm, R. C. Olson, P. S. Schnare, Mark Yu, and the proposer.

Collinear Points on a Graph

5810 [1971, 798]. *Proposed by Simeon Reich, Israel Institute of Technology, Haifa*

Let $f(x)$ be continuous on $[a,b]$ and differentiable at a and b. If $f'(a) = f'(b)$, then there is a number $H > 0$ such that corresponding to any h, $0 < h \leqq H$, there exists d, $a < d < b - h$, such that $[f(d+h) - f(d)]/h = [f(d) - f(a)]/(d-a)$.

Solution by Wayne Roberts, Macalester College, and Dale Varberg, Hamline University. We may assume that $a = 0 = f(a)$. Let $g(x) = f(x)/x$ with $g(0) = g(0+) = f'(0)$ and note that $g'(b) = -[(f(b)/b) - f'(b)]/b = -[g(b) - g(0)]/b$. From this expression it is clear that if $g(b) > g(0)$, then $g'(b) < 0$, so the maximum of g on $[0,b]$ cannot occur at b, and since $g(b) > g(0)$, it cannot occur at 0. Similar considerations for the case $g(b) < g(0)$ and the standard argument for $g(b) = g(0)$ quickly establish that g, a continuous function, achieves either its maximum or its

minimum at an interior point p of $[0,b]$. Let $H < \min(p, b-p)$. Then, for fixed $h \in (0,H]$, $k(x) = [g(x+h) - g(x)]/h$ must have opposite signs at $p-h$ and p, and being continuous must be 0 at some point $d \in [p-h,p]$. But

$$0 = \frac{g(d+h)-g(d)}{h} = \frac{1}{h}\left[\frac{f(d+h)}{d+h} - \frac{f(d)}{d}\right].$$

from which the conclusion follows.

Also solved by D. Borwein, Mats Broberg, D. O. Davies (England), Gary Gundersen, D. A. Hejhal, Terjéki József (Hungary), O. P. Lossers (Netherlands), G. C. Schmidt, and the proposer.

Collinear Points on a Monotonic Polygon

5811 [1971, 798]. *Proposed by T. C. Brown, Simon Fraser University*

Let S be a nonempty subset of the plane such that for each x in S exactly one of $x + (0,1)$ and $x + (1,0)$ also belongs to S. Prove or disprove that for each positive integer k there is a line in the plane (perhaps different lines for different k) which contains at least k points of S.

Solution by P. L. Montgomery, San Rafael, California. We shall prove the assertion. Assume $(0,0) \in S$. We define $f(n)$, $g(n)$ by: $f(0) = g(0) = 0$; if n is such that $(f(n), g(n)) \in S$, either $(f(n)+1, g(n)) \in S$ or $(f(n), g(n)+1) \in S$. In the former case let $f(n+1) = f(n)+1$ and $g(n+1) = g(n)$; in the latter, let $f(n+1) = f(n)$ and $g(n+1) = g(n)+1$. In either case $(f(n+1), g(n+1)) \in S$. By induction, $f(n) + g(n) = n$ for all n.

Since $0 \leqq (f(n)/n) \leqq 1$ for all n, the Bolzano-Weierstrass Theorem implies $\{f(n)/n\}$ has a limit point L. If $f(n)/n$ takes on the same value infinitely many times, we have a rational number x for which $f(n) = xn$ and $g(n) = (1-x)n$ for infinitely many n. Otherwise let a/b be a rational number for which $b > k$ and

$$\left|L - \frac{a}{b}\right| < \frac{1}{2b^2}.$$

Let h be the integer-valued function defined by $h(n) = bf(n) - an$. If m is a positive integer, let $N = (2m+1)(k-1)$. Then, for $0 \leqq n \leqq N$, either k of the $N+1$ values $h(n)$ are equal or one lies outside the interval $[-m, m]$. In the former case we are done; in the latter case

$$\left|\frac{h(n)}{n}\right| \geqq \frac{m+1}{N} > \frac{1}{2k} \tag{1}$$

for some n. By making m arbitrarily large we get infinitely many n satisfying (1).

For any such n

$$\left|\frac{f(n)}{n}-\frac{a}{b}\right|>\frac{1}{2kb}.$$

Letting

$$\varepsilon = \frac{1}{2kb}-\frac{1}{2b^2}>0 \quad \text{implies}$$

$$\left|\frac{f(n)}{n}-L\right|>\varepsilon \quad \text{for infinitely many } n:$$

say $\{f(n)/n\}-L>\varepsilon$ for infinitely many n.

Let $r=p/q$ be a rational number for which

$$L+\tfrac{1}{2}\varepsilon<r<L+\varepsilon.$$

Recalling that L is a limit point of $\{f(n)/n\}$, we get $f(n)>rn$ for infinitely many n and $f(n)<rn$ for infinitely many n. (If, instead, $f(n)/n-L<-\varepsilon$ for infinitely many n, let $L-\varepsilon<r<L-\frac{1}{2}\varepsilon$.)

Let $F(n)=qf(n)-pn$ for each n. Then $|F(n+1)-F(n)|\leqq q$ for all n. Since F changes sign infinitely often, F must take on some value F_0 infinitely many times; for this F_0, the set S contains infinitely many points on the line $qf(n)-pf(n)-pg(n)=F_0$.

Also solved by Robert Breusch, Mats Broberg (Sweden), K. A. Brons, Don Coppersmith, J. W. Hardy, Jr., Jacques Justin (France), Ivan Korec (Czechoslovakia), and the proposer.

An 'Elliptic' Integral

5812 [1971, 798]. *Proposed by Paul Monsky, Brandeis University*

If $f=x^4+4x^3-6x^2+4x+1$, evaluate

$$I=\int\frac{x\,dx}{\sqrt{f}}.$$

Solution by Leonard Carlitz, Duke University. Let $R(x)=x^4+ax^3+bx^2+cx+d$, where a, b, c, d are rational, be a quartic without repeated factors. Since the integral

$$\int\frac{dx}{\sqrt{R(x)}}$$

is not elementary, there is at most one value of the constant k such that

$$(*)\qquad \int\frac{(x+k)dx}{\sqrt{R(x)}}$$

is elementary. M. P. Tchebyshef (Journal de Mathématique (2), vol. 9 (1864), pp. 242–246) has shown how one can determine, in a non-tentative manner, when (*) is elementary. The method consists of expanding $\sqrt{R(x)}$ into a continued fraction of the type

$$a_0(x) + \frac{1}{a_1(x) +}\, \frac{1}{a_2(x) +} \cdots,$$

where the $a_n(x)$ are polynomials in x. If none of the polynomials $a_1(x), a_2(x), \cdots$ is of the second degree, then (*) is not elementary; actually it suffices to examine only the first N of these polynomials, where N is explicitly defined. If, on the other hand, $a_n(x)$ is the first denominator of the second degree, then the integral (*) is equal to

$$\frac{1}{2\lambda} \log \frac{\phi(x) + \sqrt{R(x)}}{\phi(x) - \sqrt{R(x)}},$$

where

$$\phi(x) = a_0(x) + \frac{1}{a_1(x) +}\, \frac{1}{a_2(x) +} \cdots \frac{1}{a_{n-1}(x)},$$

and λ is a certain positive integer.

To apply this method to the integral I where $f(x) = (x+1)^4 - 12x^2$, it is convenient to put $R(x) = x^4 - 12(x-1)^2$, so that $f(x) = R(x+1)$.

It can be verified that

$$\sqrt{R(x)} = x^2 - 6 + \cfrac{12}{x + 2 + \cfrac{-2}{x + 1 + \cfrac{-3}{x + 1 + \cfrac{-2}{x + 2 + \cfrac{6}{x^2 - 6 + \cdots}}}}}.$$

Then

$$\phi(x) = x^2 - 6 + \cfrac{12}{x + 2 + \cfrac{-2}{x + 1 + \cfrac{-3}{x + 1 + \cfrac{-2}{x + 2}}}}.$$

We find that $\phi(x) = A(x)/B(x)$, where

$$A(x) = x^6 + 6x^5 - 36x^3 + 72x, \qquad B(x) = x^4 + 6x^3 + 6x^2 - 12x - 12.$$

By the general theory

$$\int \frac{(x+k)dx}{\sqrt{R(x)}} = \frac{1}{12}\log\frac{A(x)+B(x)\sqrt{R(x)}}{A(x)-B(x)\sqrt{R(x)}}$$
$$= \frac{1}{12}\log\frac{(A(x)+B(x)\sqrt{R(x)})^2}{A^2(x)-B^2(x)R(x)}.$$

By direct computation, $A^2(x)-B^2(x)R(x) = 12^3$. Also it is easily verified that $A'(x) = 6(x-1)B(x)$. Combining these equations, we get $\frac{1}{2}B(x)R'(x)+B'(x)R(x) = 6(x-1)A(x)$. It follows that

$$\frac{d}{dx}\log(A(x)+B(x)\sqrt{R(x)}) = \frac{6(x-1)}{\sqrt{R(x)}}$$

or, what is the same thing,

$$\int \frac{(x-1)dx}{\sqrt{R(x)}} = \frac{1}{6}\log(A(x)+B(x)\sqrt{R(x)}).$$

Finally, therefore,

$$\int \frac{x\,dx}{\sqrt{f(x)}} = \frac{1}{6}\log(A(x+1)+B(x+1)\sqrt{f(x)}).$$

Also solved by V. S. Blanco, C. A. Bridger, Fred Dodd, Václav Konečný, and the proposer.

Commutator of Operators on a Hilbert Space

5813 [1971, 798]. *Proposed by A. R. Barron, Brandeis University*

Show that if A is a bounded operator and B is a self adjoint operator, on some Hilbert space, then

$$[B,[B,A]] = 0 \quad \text{implies} \quad [B,A] = 0.$$

Note: $[X,Y] = XY - YX$.

Solution by Joel Anderson, California Institute of Technology. This is an easy consequence of the Kleinecke-Shirokov theorem (*A Hilbert Space Problem Book*, P. R. Halmos, problem 184): If P and Q are bounded operators and $R = PQ - QP$ commutes with P, then R is quasinilpotent. The hypothesis $[B,[B,A]] = 0$ implies that $C = B(A-A^*) - (A-A^*)B$ commutes with B. Hence, C is quasinilpotent. But C is self-adjoint, hence $C = 0$. Similarly $iB(A+A^*) - i(A+A^*)B = 0$ and it follows that $BA - AB = 0$.

This fact has been known at least since 1959 (see S. Sakai, *On some problems of C^*-algebras*, Tôhoku Math. J. (2) 11(1959), 453–455, Theorem 1). In fact it is shown (the solver's Ph.D. thesis, Indiana University, 1971) that if T is a bounded

operator which commutes with B, and X is any bounded operator, then

$$\| T - [B,X] \| \geqq \| T \| .$$

(See the solver's forthcoming paper, *On normal derivations.*)

Also solved by Cecilia H. Brook, S. L. Campbell, J. A. Deddens, Ellen Hertz, A. A. Jagers (Netherlands), E. M. Klein, J. S. Lancaster, Kazuhiro Tamaki (Japan), Olga Taussky Todd, J. P. Williams, and the proposer.

Editor's Note. Olga Taussky Todd has written that the result of the problem in a more general form is contained in a paper by C. R. Putnam, *On normal operators in Hilbert Space*, Amer. J. Math. 73 (1951), 357–362. The result was used by Professor Todd and T. Køto in *Commutators of A and A**, J. Washington Acad. Sci. 46, Feb. 1956, 38–39, Theorem 5.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

Topics in Modern Mathematics for Teachers. By Anthony L. Peressini and Donald R. Sherbert. Holt, Rinehart & Winston, New York, 1971. xi+434 pp. $11.95. (Telegraphic Review, January 1972.)

The authors have presented short but readable accounts of a number of diverse topics. Their success at retaining a certain amount of depth with brevity is commendable. Any teacher using the book can pick and choose which topics he wants to cover because of their relative independence. The authors have paid a penalty for this independence in terms of the book's uneven quality. Chapter 7, on graph theory, is intriguing reading, while Chapter 6, on an equally interesting topic, Boolean algebra, is dull, save for the last section. Each chapter includes a section relating its content to the school curriculum. These sections contain excellent references, but the relationships drawn to school curricula are trivial.

The authors have acknowledged two different motivations in selecting topics:

INDEX TO VOLUME 79, 1972

THE AMERICAN MATHEMATICAL MONTHLY

AUTHOR INDEX

KEY WORDS AND PHRASES INDEX

PROBLEMS AND SOLUTIONS

PROBLEMS PROPOSED

PROBLEMS SOLVED

SOLUTIONS

Numbers in **boldface** type refer to problems, those in lightface, to pages

REVIEWS

Abbrevations: (TR)—Telegraphic Review; (NP)—Notable Paper.
Names of authors are in ordinary type, those of reviewers in capitals.

FILMS

Acknowledgment. The following have generously helped in evaluating books: FREEMAN DYSON, REUBEN HERSH, R. KALMAN, NATHANIEL MACON, R. MCDOWELL, GEORGE POLYA, SAMUEL SCHECHTER, JULIAN WEISSGLASS.

TELEGRAPHIC REVIEWS

NEWS AND NOTICES

PERSONAL ITEMS

GENERAL INFORMATION

NECROLOGY